



TUGAS AKHIR - SS141501

**PEMODELAN DAN PEMETAAN KEJADIAN DIARE
DI KOTA SURABAYA MENGGUNAKAN
*GEOGRAPHICALLY WEIGHTED NEGATIVE BINOMIAL
REGRESSION***

**FITRI ERNAWATI
NRP 1313 105 029**

**Dosen Pembimbing
Ir. Mutiah Salamah, M.Kes**

**Program Studi S1 Statistika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Institut Teknologi Sepuluh Nopember
Surabaya 2015**



FINAL PROJECT - SS141501

MODELLING AND MAPPING OF THE INCIDENCE DIARE IN SURABAYA USING *GEOGRAPHICALLY WEIGHTED NEGATIVE BINOMIAL REGRESSION*

FITRI ERNAWATI
NRP 1313 105 029

Supervisor
Ir. Mutiah Salamah, M.Kes

Undergraduate Programme of Statistics
Faculty of Mathematics and Natural Sciences
Institut Teknologi Sepuluh Nopember
Surabaya 2015

LEMBAR PENGESAHAN

PEMODELAN DAN PEMETAAN KEJADIAN DIARE DI KOTA SURABAYA MENGGUNAKAN *GEOGRAPHICALLY WEIGHTED NEGATIVE BINOMIAL REGRESSION*

TUGAS AKHIR

Diajukan Untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Sains
pada
Program Studi S-1 Jurusan Statistika
Fakultas Matematika Dan Ilmu Pengetahuan Alam
Institut Teknologi Sepuluh Nopember

Oleh:

FITRI ERNAWATI

NRP 1313 105 029

Disetujui oleh Pembimbing Tugas Akhir :

Ir. Mutiah Salamah, M.Kes

NIP. 19571007 198303 2 001



Mengetahui,

Ketua Jurusan Statistika FMIPA - ITS



Dr. Muhammad Mashuri, MT
NIP. 19620408 198701 1 001

SURABAYA, JULI 2015

PEMODELAN DAN PEMETAAN KEJADIAN DIARE DI KOTA SURABAYA MENGGUNAKAN GEOGRAPHICALLY WEIGHTED NEGATIVE BINOMIAL REGRESSION

Nama Mahasiswa : Fitri Ernawati
NRP : 1313 105 029
Program Studi : Sarjana
Dosen Pembimbing : Ir. Mutiah Salamah, M.Kes

Abstrak

Surabaya adalah ibu kota Provinsi Jawa Timur yang sekaligus menjadi kota terbesar kedua di Indonesia. Selain itu Kota Surabaya juga terkenal dengan kepadatan penduduknya yang menyebabkan di beberapa wilayah Kota Surabaya terlihat kumuh dan buruk akan sanitasi lingkungan. Lingkungan dengan sanitasi kesehatan yang buruk berpotensi menimbulkan adanya bakteri penyebab penyakit, salah satunya adalah diare. Penyakit diare adalah keadaan dimana penderita mengalami rangsangan buang air besar terus menerus dan tinja memiliki kandungan air berlebihan. Kejadian diare tertinggi di Provinsi Jawa Timur yaitu Kota Surabaya sebesar 95.105 kejadian pada tahun 2013. Berdasarkan hal tersebut dibutuhkan cara untuk mengurangi kejadian diare yaitu dengan mengetahui faktor-faktor apa saja yang mempengaruhinya, sehingga dalam penelitian ini dilakukan pemodelan dan pemetaan kejadian diare di Kota Surabaya menggunakan geographically weighted negative binomial regression (GWNBR). GWNBR adalah salah satu metode statistika yang digunakan untuk mengatasi overdispersi dan memperhitungkan faktor spasial sebagai variabel bebas. Hasil penelitian menunjukkan bahwa terdapat 6 kelompok kecamatan di Kota Surabaya berdasarkan kesamaan variabel yang berpengaruh terhadap kejadian diare. Variabel yang berpengaruh terhadap kejadian diare di semua kecamatan Kota Surabaya adalah kepadatan penduduk, rasio sarana kesehatan puskesmas dan rasio dokter umum.

Kata Kunci : GWNBR, Kejadian Diare, Kepadatan Penduduk

“Halaman ini sengaja dikosongkan”

MODELLING AND MAPPING OF THE INCIDENCE DIARE IN SURABAYA USING GEOGRAPHICALLY WEIGHTED NEGATIVE BINOMIAL REGRESSION

Name : Fitri Ernawati
NRP : 1313 105 029
Department : Statistics
Supervisor : Ir. Mutiah Salamah, M.Kes

Abstract

Surabaya is the capital of East Java province which is also the second largest city in Indonesia. In addition, Surabaya is also famous for density of population is causing in some areas Surabaya looks rundown and bad environmental sanitation. Environmental sanitation in poor healthy with potential presence bacteria of disease, one of which is diare. Diare is disease a condition where patients experienced bowel stimulation of continuous and excessive fecal water content. The highest incidence of diare with the conditions of high population density are also in the province of East Java, Surabaya amounting to 95.105 events in 2013. Therefore, in this research modelling and mapping of the incidence diare in Surabaya using geographically weighted negative binomial regression (GWNBR). GWNBR is one of the statistical methods used to overcome overdispersion and take into account the spatial factors as independent variables. The results showed that there are six groups of districts in the city of Surabaya based on common variables that affect the incidence of diare. Variables that influence the incidence of diare in all districts of Surabaya is the population density, the ratio of health facilities and ratio of general practitioners.

Keywords : GWNBR, Incidence Diare, Density Population

“Halaman ini sengaja dikosongkan”

KATA PENGANTAR

Segala puji dan syukur dihaturkan kehadiran Allah SWT atas segala rahmat, hidayah, nikmat, dan petunjuk-Nya serta tak lupa Sholawat dan salam kepada Rasulullah SAW sehingga penulis bisa menyelesaikan Tugas Akhir yang berjudul

“Pemodelan dan Pemetaan Kejadian Diare di Kota Surabaya Menggunakan *Geographically Weighted Negative Binomial Regression*”

Selama proses penyusunan hingga penyelesaian Tugas Akhir ini tidak lepas dari doa, saran, bimbingan serta dukungan dari berbagai pihak. Oleh karena itu, penulis ingin mengucapkan terima kasih kepada :

1. Ibu Ir. Mutiah Salamah, M.Kes selaku dosen pembimbing atas ilmu, doa, motivasi, kesabaran, serta waktu dan perhatian selama membimbing hingga terselesainya Tugas Akhir ini dengan baik.
2. Bapak Dr. Purhadi, M.Sc dan Ibu Shofi Andari, M.Si selaku dosen penguji atas ilmu dan saran yang telah diberikan untuk Tugas Akhir yang lebih baik.
3. Bapak Dr. Muhammad Mashuri, MT selaku Ketua Jurusan Statistika FMIPA ITS.
4. Bapak Prof. Dr. I nyoman Budiantara, M.Si selaku dosen wali yang telah memberikan motivasi untuk mendapatkan nilai terbaik selama masa perkuliahan.
5. Ibu Dra. Lucia Aridinanti, MT selaku Ketua Program Studi S1 Statistika ITS.
6. Keluarga Besar Jurusan Statistika FMIPA ITS untuk segala dukungan dan bantuan selama penulis menempuh studi di Lintas Jalur S1 Statistika ITS.
7. Pihak Dinas Kesehatan Kota Surabaya yang telah memberikan kemudahan dalam proses permintaan data.

8. Orang tua dan keluarga yang selalu memberi dukungan baik materi maupun doa.
9. Sahabat lama Rully, Cicik, Ria, Kuncah, Ani yang selalu menghibur dan memberi semangat penulis.
10. Sahabat terbaik Fifi, Novil, Ida, Silvi, Vellin, Arinda, Mirna, Uyun, Wardah, Tanti, Meita, Woro, Diyah, Evi, Alma, Sinta yang selalu menemani cerita saat suka dan duka penulis.
11. Tim GW Fitria dan Urifah, serta tim RBS Fitrah, Ega dan Alfi yang telah menemani diskusi dan kegalauan metode.
12. Teman-teman LJ Statistika ITS 2013 untuk motivasi, canda tawa, dan kebersamaan selama ini.
13. Semua pihak yang telah membantu kelancaran penyelesaian Tugas Akhir ini dalam segala bentuk yang tidak dapat disebutkan satu per satu.

Penulis menyadari bahwa Tugas Akhir ini masih jauh dari sempurna. Oleh sebab itu, diharapkan saran dan kritik yang membangun dari berbagai pihak untuk dijadikan bahan pertimbangan dalam penyusunan laporan berikutnya.

Surabaya, Juli 2015

Penulis

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
LEMBAR PENGESAHAN	v
ABSTRAK	vii
ABSTRACT	ix
KATA PENGANTAR	xi
DAFTAR ISI	xiii
DAFTAR GAMBAR	xv
DAFTAR TABEL	xvii
DAFTAR LAMPIRAN	xix
BAB I PENDAHULUAN	
1.1 Latar Belakang	1
1.2 Perumusan Masalah	4
1.3 Tujuan Penelitian	4
1.4 Manfaat Penelitian	4
1.5 Batasan Masalah	4
BAB II TINJAUAN PUSTAKA	
2.1 Statistika Deskriptif.....	5
2.2 Multikolinieritas	5
2.3 Regresi Poisson.....	6
2.3.1 Estimasi Parameter Regresi Poisson.....	7
2.3.2 Pengujian Parameter Regresi Poisson	9
2.4 Overdispersi Regresi Poisson.....	10
2.5 Regresi Binomial Negatif	11
2.5.1 Estimasi Parameter Regresi Binomial Negatif.....	12
2.5.2 Pengujian Parameter Regresi Binomial Negatif	15
2.6 Pengujian Aspek Data Spasial	16
2.6.1 Pengujian Dependensi Spasial	16
2.6.2 Pengujian Heterogenitas Spasial	17
2.7 Model GWNBR	18
2.7.1 Estimasi Parameter Model GWNBR	20
2.7.2 Pengujian Parameter Model GWNBR	27
2.7.3 Penentuan Bandwidth dan Pembobot Optimum	28

2.8 Penyakit Diare	29
2.9 Faktor-Faktor Mempengaruhi Penyakit Diare	29
BAB III METODOLOGI PENELITIAN	
3.1 Sumber Data	33
3.2 Variabel Penelitian	33
3.3 Langkah Analisis	35
BAB IV ANALISIS DAN PEMBAHASAN	
4.1 Deskripsi Jumlah Kejadian Diare Beserta Faktor- Faktor yang diduga Mempengaruhi	39
4.1.1 Jumlah Kejadian Diare Tahun 2013 di Kota Surabaya	40
4.1.2 Persentase Rumah Sehat	42
4.1.3 Persentase RT Memiliki Jamban Sehat	43
4.1.4 Persentase RT Memiliki Tempat Sampah Sehat	44
4.1.5 Persentase RT Miskin	45
4.1.6 Persentase RT RT berPHBS	46
4.1.7 Kepadatan Penduduk	47
4.1.8 Rasio Sarana Kesehatan Puskesmas	48
4.1.9 Rasio Dokter Umum	49
4.2 Pemodelan Jumlah Kejadian Diare	50
4.2.1 Pemeriksaan Multikolinieritas	50
4.2.2 Pemodelan Regresi Poisson dan Regresi Binomial Negatif	51
4.2.3 Pemodelan Jumlah Kejadian Diare di Kota Surabaya Menggunakan Metode GWNBR	54
4.2.3.1 Pengujian Dependensi dan Heterogenitas Spasial	54
4.2.3.2 Pengujian Signifikansi Parameter GWNBR	56
4.2.3.3 Pemilihan Model Terbaik	62
BAB V KESIMPULAN DAN SARAN	
5.1 Kesimpulan	63
5.2 Saran	64
DAFTAR PUSTAKA	65
LAMPIRAN	67

DAFTAR TABEL

Tabel 3.1	Variabel Penelitian	33
Tabel 3.2	Struktur Data Penelitian	35
Tabel 4.1	Karakteristik Jumlah Kejadian Diare Beserta Faktor-Faktor yang diduga Mempengaruhinya	39
Tabel 4.2	Koefisien Korelasi Antar Variabel Prediktor	50
Tabel 4.3	Nilai VIF Masing-Masing Variabel Prediktor	51
Tabel 4.4	Penaksiran Parameter Model Regresi Poisson	51
Tabel 4.5	Penaksiran Parameter Model Regresi Binomial Negatif	53
Tabel 4.6	Penaksiran Parameter Model GWNBR	57
Tabel 4.7	Pengelompokkan Kecamatan Berdasarkan Variabel Signifikan dalam Model GWNBR.....	58
Tabel 4.8	Parameter Model GWNBR di Kecamatan Kenjeran	60
Tabel 4.9	Hasil Perhitungan Kriteria AIC	62

“Halaman ini sengaja dikosongkan”

DAFTAR GAMBAR

Gambar 3.1	Diagram Alir Langkah Analisis	37
Gambar 4.1	Persebaran Jumlah Kejadian Diare Kota Surabaya	41
Gambar 4.2	Persebaran Persentase Rumah Sehat	42
Gambar 4.3	Persebaran Persentase RT Memiliki Jamban Sehat	43
Gambar 4.4	Persebaran Persentase RT Memiliki Tempat Sampah Sehat.....	44
Gambar 4.5	Persebaran Persentase RT Miskin.....	45
Gambar 4.6	Persebaran Persentase RT ber-PHBS.....	46
Gambar 4.7	Persebaran Kepadatan Penduduk	47
Gambar 4.8	Persebaran Rasio Sarana Kesehatan Puskesmas	48
Gambar 4.9	Persebaran Rasio Dokter Umum.....	49
Gambar 4.10	Pemetaan Kecamatan Berdasarkan Faktor-Faktor Mempengaruhi Kejadian Diare di Kota Surabaya	59

“Halaman ini sengaja dikosongkan”

DAFTAR LAMPIRAN

Lampiran 1	Data Jumlah Kejadian Diare Beserta Faktor-Faktor yang diduga Mempengaruhinya.....	67
Lampiran 2	Statistika Deskriptif Jumlah Kejadian Diare Beserta Faktor-Faktor yang diduga Mempengaruhinya.....	69
Lampiran 3	Hasil Pemeriksaan Multikolinieritas	69
Lampiran 4	Hasil Analisis Regresi Poisson.....	70
Lampiran 5	Hasil Analisis Regresi Binomial Negatif	71
Lampiran 6	Hasil Analisis Pengujian Aspek Data Spasial	72
Lampiran 7	Jarak Antar Kecamatan	73
Lampiran 8	Matriks Pembobot Geografis	74
Lampiran 9	Z Hitung Pengujian Parsial GWNBR	75
Lampiran 10	Koefisien Parameter	76
Lampiran 11	<i>Syntax</i> Analisis Menggunakan <i>Software R</i>	77
Lampiran 12	Prediksi Kejadian Diare Berdasarkan Model GWNBR.....	82

“Halaman ini sengaja dikosongkan”

BAB I

PENDAHULUAN

1.1 Latar Belakang

Menurut WHO secara klinis penyakit diare adalah penyakit bertambahnya defaksi (buang air besar) lebih dari biasanya (umumnya lebih dari tiga kali dalam sehari) disertai dengan perubahan konsisten tinja menjadi cair dengan atau tanpa darah. Diare adalah sebuah penyakit dimana penderita mengalami rangsangan buang air besar terus menerus dan tinja memiliki kandungan air berlebihan (Wikipedia, 2015). Penyakit ini dapat digolongkan penyakit ringan, tetapi jika terjadi secara mendadak dan tidak mendapatkan perawatan yang tepat maka diare dapat berakibat fatal misalnya dehidrasi berat yang bisa menyebabkan kematian. Diare dapat menular melalui kuman, bakteri dan virus. Selain itu kondisi iklim yang tidak menentu, kebersihan lingkungan seperti terjadi banjir, sanitasi yang buruk dan juga kurangnya kesadaran masyarakat terhadap kesehatan turut berperan dalam peningkatan kasus diare. Dalam kondisi hidup yang bersih dan dengan makanan yang mencukupi serta air tersedia, pasien biasanya sembuh dari infeksi virus umum dalam beberapa hari dan paling lama satu minggu. Namun untuk pasien yang kurang gizi, diare dapat mengancam jiwa bila tanpa perawatan yang tepat.

Berdasarkan hasil penelitian Purwaningsih (2009) mengenai analisis hubungan antara kondisi sanitasi, air bersih, dan penderita diare di Jawa Timur menunjukkan bahwa faktor-faktor yang berpengaruh terhadap penderita diare adalah sumber air minum terlindung, penggunaan jamban sendiri dan jumlah anggota rumah tangga. Begitu juga menurut Arumsari (2011) mengenai pemodelan penderita diare di Kabupaten Sumedang adalah jarak fasilitas kesehatan dan fasilitas air minum. Sedangkan menurut Andjarsari (2012) mengenai analisis faktor-faktor yang mempengaruhi terjangkitnya penyakit diare menggunakan regresi logistik biner menunjukkan bahwa jenis

kelamin, pendidikan, usia, pekerjaan, kualitas fisik air minum keruh, tempat penampungan air sebelum dimasak, sebelum makan mencuci tangan pakai sabun dan tempat pembuangan air besar merupakan faktor yang mempengaruhi terjangkitnya penyakit diare.

Capaian penemuan diare cenderung mengalami peningkatan dalam lima tahun terakhir yaitu dari tahun 2009 hingga tahun 2013 di Provinsi Jawa Timur dan jumlah kasus diare yang paling tinggi adalah Kota Surabaya dengan jumlah penduduk yang paling tinggi pula yaitu sebanyak 95.105 kejadian diare (Laporan Profil Kesehatan Provinsi Jawa Timur, 2013). Kota Surabaya merupakan ibu kota Provinsi Jawa Timur dan tergolong dalam kategori kota besar serta kepadatan penduduknya. Oleh karena itu dalam perkembangannya Kota Surabaya selalu melakukan upaya pembangunan untuk menjadi yang terbaik dalam segala bidang. Bidang kesehatan adalah salah satu hal utama yang dilakukan untuk peningkatan menjadi lebih baik demi menunjang pembangunan sistem dan infrastruktur Kota Surabaya. Sekilas Kota Surabaya merupakan kota yang indah tanpa adanya masalah, namun ternyata ada daerah tertentu yang butuh sorotan dan bantuan upaya pemerintah terutama dalam hal sanitasi lingkungan untuk kesehatan.

Penyakit diare dapat ditularkan melalui kondisi lingkungan sekitar, misalnya dengan keadaan kondisi lingkungan yang kumuh. Lingkungan yang kumuh tersebut dampak akibat adanya kepadatan penduduk. Kepadatan penduduk adalah kondisi di mana bertambahnya jumlah penduduk, pertambahan penduduk tersebut dapat mengakibatkan keterbatasan akan lahan khususnya lahan untuk pemukiman, penurunan akibat limbah kota, peningkatan pencemaran udara akibat pemanfaatan transportasi kota dan meningkatnya angka kemiskinan. Hal tersebut menjadi pemicu penurunan kualitas kota, salah satunya dalam bidang kesehatan. Kesehatan merupakan indeks pembangunan manusia yang menjadi salah satu indikator pembangunan kesehatan. Pembangunan kesehatan harus dapat menjangkau kelompok-

kelompok penduduk yang beresiko tinggi sebagai penyumbang kejadian sakit dan kematian. Oleh karena itu dengan semakin terbatasnya lahan dengan kepadatan penduduk maka penyebaran akan kuman, virus dan bakteri dari penyakit menular akan lebih mudah melalui kondisi lingkungan sekitar.

Berdasarkan kondisi kepadatan penduduk di Kota Surabaya dengan kejadian diare yang tinggi maka dalam penelitian ini dilakukan analisis mengenai hal tersebut dengan memperhatikan faktor spasial yang ditimbulkan dari kepadatan penduduk masing-masing wilayah di Kota Surabaya. Salah satu metode statistika yang digunakan untuk mengatasi overdispersi dan memperhitungkan faktor spasial sebagai variabel bebas yang dapat mempengaruhi variabel respon adalah *geographically weighted negative binomial regression* (GWNBR). Sehingga didapatkan hasil bahwa setiap wilayah mempunyai perbedaan karakteristik satu sama lain sehingga menghasilkan model yang juga berbeda.

Penelitian dengan menggunakan metode GWNBR yang telah dilakukan diantaranya Nandasari (2014) yaitu untuk memodelkan jumlah kejadian luar biasa difteri di Jawa Timur Pratama (2015) untuk memodelkan jumlah kasus penyakit *Tuberculosis* (TBC) di Provinsi Jawa Barat serta Juniardi LC dan Salamah M. (2015) melakukan analisis faktor-faktor yang mempengaruhi jumlah kasus Kusta di Jawa Timur pada Tahun 2013 Menggunakan *Geographically Weighted Negative Binomial Regression* (GWNBR). Oleh karena itu pada penelitian ini dilakukan pemodelan kasus diare berdasarkan faktor-faktor yang mempengaruhi pada kecamatan di Kota Surabaya dengan menggunakan GWNBR. Metode tersebut diharapkan dapat lebih menjelaskan dan mengidentifikasi faktor-faktor yang mempengaruhi terjadinya kejadian diare pada kecamatan di Kota Surabaya dengan menyesuaikan karakteristik masing-masing wilayah (faktor spasial).

1.2 Perumusan Masalah

Kota Surabaya terkenal dengan kepadatan penduduknya yang menyebabkan di beberapa wilayah Kota Surabaya terlihat kumuh dan buruk akan sanitasi lingkungan. Lingkungan dengan sanitasi kesehatan yang buruk berpotensi menimbulkan adanya bakteri penyebab penyakit, salah satunya adalah diare. Kejadian diare tertinggi di Provinsi Jawa Timur yaitu Kota Surabaya sebesar 95.105 kejadian pada tahun 2013. Berdasarkan hal tersebut perlu dilakukan upaya untuk mengurangi kejadian diare yaitu dengan mengetahui faktor-faktor yang mempengaruhinya, sehingga dalam penelitian ini dilakukan pemodelan dan pemetaan kejadian diare berdasarkan faktor-faktor yang mempengaruhi di Kota Surabaya menggunakan *geographically weighted negative binomial regression* (GWNBR). GWNBR adalah salah satu metode statistika yang digunakan untuk mengatasi overdispersi dan memperhitungkan faktor spasial sebagai variabel bebas.

1.3 Tujuan

Berdasarkan perumusan masalah diatas maka tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut.

1. Mengetahui karakteristik kejadian diare berdasarkan faktor-faktor yang diduga mempengaruhi di kecamatan Kota Surabaya.
2. Menganalisis dan memodelkan kejadian diare di kecamatan Kota Surabaya berdasarkan faktor-faktor yang mempengaruhi menggunakan metode GWNBR.

1.4 Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah dapat menerapkan aplikasi ilmu statistika khususnya dalam bidang kesehatan. Selain itu dapat mengetahui faktor-faktor apa saja yang mempengaruhi terjadinya penyakit diare di Kota Surabaya. Manfaat lain juga dapat membantu memberikan informasi kepada Dinas Kesehatan serta Pemerintah Kota Surabaya untuk solusi dalam meningkatkan pembangunan khususnya bidang kesehatan.

BAB II TINJAUAN PUSTAKA

2.1 Statistika Deskriptif

Statistika deskriptif adalah metode-metode yang berkaitan dengan pengumpulan dan penyajian suatu gugus data sehingga memberikan informasi yang berguna (Walpole, 1995). Statistika deskriptif membahas pengumpulan data, meringkas dan menyajikan data, serta melakukan pengukuran pemusatan dan penyebaran untuk memperoleh informasi yang lebih menarik, berguna dan lebih mudah dipahami. Salah satu bentuk penyajian statistika deskriptif adalah dengan menggunakan peta tematik. Peta tematik adalah gambaran dari sebagian permukaan bumi yang dilengkapi dengan informasi tertentu, baik di atas maupun di bawah permukaan bumi yang mengandung tema tertentu, baik data kualitatif maupun data kuantitatif. Peta tematik sangat erat kaitannya dengan sistem informasi geografis (SIG) karena pada umumnya *output* dari SIG adalah berupa peta tematik. Baik yang berbentuk digital maupun masih berbentuk peta kertas (Barus dan Wiradisastra, 2000).

2.2 Multikolinieritas

Salah satu syarat yang harus dipenuhi dalam pembentukan model regresi dengan beberapa variabel prediktor adalah tidak ada kasus multikolinieritas atau tidak terdapat korelasi antara satu variabel prediktor dengan variabel prediktor yang lain. Dalam model regresi, adanya korelasi antar variabel prediktor menyebabkan taksiran parameter regresi yang dihasilkan akan memiliki eror yang sangat besar. Menurut Hocking (1996) pendeteksian kasus multikolinieritas dilakukan menggunakan kriteria nilai *variance inflation factor* (VIF). Jika nilai VIF lebih besar dari 10 menunjukkan adanya multikolinieritas antarvariabel prediktor. Nilai VIF dinyatakan sebagai berikut.

$$VIF = \frac{1}{1 - R_j^2} \quad ; j = 1, 2, \dots, p \quad (2.1)$$

dengan R_j^2 adalah koefisien determinasi antara satu variabel prediktor dengan variabel prediktor lainnya. Solusi untuk mengatasi adanya kasus multikolinieritas yaitu dengan mengeluarkan variabel prediktor satu per satu mulai dari yang memiliki nilai VIF paling besar.

2.3 Regresi Poisson

Regresi Poisson merupakan model regresi nonlinier yang digunakan untuk data jumlah (*count*) dimana variabel respon mengikuti distribusi Poisson (Agresti, 2002). Menurut Walpole (1995), distribusi Poisson adalah distribusi probabilitas acak yang menyatakan banyaknya sukses dari suatu percobaan.

Jika variabel random diskrit (y) merupakan distribusi Poisson dengan parameter μ maka fungsi probabilitas dari distribusi Poisson dapat dinyatakan sebagai berikut.

$$f(y, \mu) = \frac{e^{-\mu} \mu^y}{y!}; y = 0, 1, 2, \dots \quad (2.2)$$

dengan μ adalah rata-rata jumlah sukses dalam variabel random y dan bilangan bulat positif ($\mu > 0$) maka $E(y) = \mu$ dan $Var(y) = \mu$

Regresi Poisson merupakan *generalized linear model* (GLM) karena terdapat 3 komponen GLM dalam regresi Poisson yaitu komponen random, komponen sistematis, dan *link function* (Agresti, 2002). Komponen random merupakan variabel respon yaitu variabel \mathbf{y} yang merupakan vektor yang terdiri dari $[y_1, y_2, \dots, y_n]^T$. Komponen sistematis yaitu vektor $\boldsymbol{\eta}$ yang terdiri dari $[\eta_1, \eta_2, \dots, \eta_n]^T$ menghubungkan dengan variabel X , sehingga bentuk umum dari $\boldsymbol{\eta} = \mathbf{X}\boldsymbol{\beta}$ di mana \mathbf{X} merupakan suatu matriks dengan elemen yang terdiri variabel prediktor, sedangkan $\boldsymbol{\beta}$ merupakan bentuk vektor dari parameter-parameter model. Masing-masing dari elemen $\boldsymbol{\eta}$ dapat dinyatakan dengan

$\eta_i = \sum_{k=0}^p \beta_k x_{ik}$ merupakan kombinasi linier dari variabel prediktor.

Komponen yang ketiga yaitu *link function* yang menghubungkan antara komponen random dan komponen sistematis. Dimisalkan $\mu = E(Y_i)$ di mana $i = 1, 2, \dots, n$. Model untuk menghubungkan μ_i dengan η_i oleh $g(\mu_i) = \eta_i$, dimana $g(\cdot)$ adalah fungsi yang diturunkan. Sehingga $g(\cdot)$ menghubungkan $E(Y_i)$ dengan variabel prediktor, dan diformulasikan sebagai berikut:

$$g(\mu_i) = \beta_0 + \sum_{k=1}^p \beta_k x_{ik} = \beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip} \quad (2.3)$$

Link function yang digunakan dalam regresi Poisson adalah \ln , sehingga $\ln(\mu_i) = \eta_i$. Dengan demikian regresi Poisson dapat dinyatakan sebagai berikut:

$$\ln(\mu_i) = \sum_{k=1}^p \beta_k x_{ik} \quad i = 1, 2, \dots, n$$

dimana:

$$\mu_i = \mu_i(\mathbf{X}_i) = \exp\left(\beta_0 + \sum_{k=1}^p \beta_k x_{ik}\right) \quad (2.4)$$

2.3.1 Estimasi Parameter Regresi Poisson

Metode yang digunakan untuk melakukan estimasi parameter regresi Poisson adalah metode *maximum likelihood estimation* (MLE). Fungsi likelihood dari regresi Poisson adalah

$$\begin{aligned} L(\boldsymbol{\beta}) &= \left(\prod_{i=1}^n \frac{\exp(-\mu_i) \mu_i^{y_i}}{y_i!} \right) \\ \ln L(\boldsymbol{\beta}) &= \ln \left(\prod_{i=1}^n \frac{\exp(-\mu_i) \mu_i^{y_i}}{y_i!} \right) \\ &= \sum_{i=1}^n \ln \left(\frac{\exp(-\mu_i) \mu_i^{y_i}}{y_i!} \right) \\ &= \sum_{i=1}^n (\ln(e^{-\mu_i}) + \ln(\mu_i^{y_i}) - \ln(y_i!)) \end{aligned}$$

$$\begin{aligned}
&= \sum_{i=1}^n (-\mu_i + y_i \ln(\mu_i) - \ln(y_i!)) \\
&= \sum_{i=1}^n \left(-e^{\mathbf{X}_i^T \boldsymbol{\beta}} + y_i \ln e^{\mathbf{X}_i^T \boldsymbol{\beta}} - \ln(y_i!) \right) \\
&= -\sum_{i=1}^n e^{\mathbf{X}_i^T \boldsymbol{\beta}} + \sum_{i=1}^n y_i \mathbf{X}_i^T \boldsymbol{\beta} - \sum_{i=1}^n \ln(y_i!) \quad (2.5)
\end{aligned}$$

Estimasi maksimum Likelihood untuk parameter β_k dinyatakan dengan $\hat{\beta}_k$ yang merupakan penyelesaian dari turunan pertama fungsi logaritma natural dari likelihood. Selanjutnya persamaan (2.5) diturunkan terhadap $\boldsymbol{\beta}^T$ dan disamakan dengan nol dan dapat diselesaikan dengan metode numerik Newton Raphson sebagai berikut:

$$\frac{\partial \ln L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}^T} = -\sum_{i=1}^n \mathbf{x}_i \exp(\mathbf{x}_i^T \boldsymbol{\beta}) + \sum_{i=1}^n y_i \mathbf{x}_i \quad (2.6)$$

Berikut ini merupakan langkah-langkah dalam estimasi parameter model regresi Poisson (Cameron & Trivedi, 1998):

1. Menentukan nilai estimasi awal parameter $\hat{\boldsymbol{\beta}}_{(0)}$

$$\hat{\boldsymbol{\beta}}_{(0)} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}$$

2. Membentuk vektor gradien g,

$$\mathbf{g}_{(p+1) \times 1}^T = \left[\frac{\partial \ln L(\boldsymbol{\beta})}{\partial \beta_0}, \frac{\partial \ln L(\boldsymbol{\beta})}{\partial \beta_1}, \dots, \frac{\partial \ln L(\boldsymbol{\beta})}{\partial \beta_p} \right]_{\beta = \hat{\beta}_{(m)}}$$

p adalah jumlah parameter yang diestimasi.

3. Membentuk matriks Hessian H:

$$\mathbf{H}(\boldsymbol{\beta}_{(m)})_{(p+1)(p+1)} = \begin{bmatrix} \frac{\partial^2 \ln L(\boldsymbol{\beta})}{\partial \beta_0^2} & \frac{\partial^2 \ln L(\boldsymbol{\beta})}{\partial \beta_0 \partial \beta_1} & \dots & \frac{\partial^2 \ln L(\boldsymbol{\beta})}{\partial \beta_0 \partial \beta_p} \\ \frac{\partial^2 \ln L(\boldsymbol{\beta})}{\partial \beta_0 \partial \beta_1} & \frac{\partial^2 \ln L(\boldsymbol{\beta})}{\partial \beta_1^2} & \dots & \frac{\partial^2 \ln L(\boldsymbol{\beta})}{\partial \beta_1 \partial \beta_p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \ln L(\boldsymbol{\beta})}{\partial \beta_0 \partial \beta_p} & \frac{\partial^2 \ln L(\boldsymbol{\beta})}{\partial \beta_1 \partial \beta_p} & \dots & \frac{\partial^2 \ln L(\boldsymbol{\beta})}{\partial \beta_p^2} \end{bmatrix}_{\beta = \hat{\beta}_{(m)}}$$

4. Mulai dari $m = 0$ dilakukan iterasi pada persamaan :

$$\hat{\beta}_{(m+1)} = \hat{\beta}_{(m)} - \mathbf{H}^{-1}(\hat{\beta}_{(m)}) \mathbf{g}(\hat{\beta}_{(m)})$$

Nilai $\hat{\beta}_{(m)}$ merupakan sekumpulan penaksir parameter yang konvergen pada iterasi ke- m .

5. Jika belum didapatkan penaksir parameter yang konvergen, maka dilanjutkan kembali langkah 4 hingga iterasi ke $m = m + 1$. Iterasi berhenti pada keadaan konvergen yaitu pada saat $\|\hat{\beta}_{(m+1)} - \hat{\beta}_{(m)}\| \leq \varepsilon$, dimana ε merupakan bilangan yang sangat kecil sekali.

2.3.2 Pengujian Parameter Regresi Poisson

Pengujian parameter digunakan untuk mengetahui pengaruh dari suatu parameter terhadap model dengan tingkat signifikansi tertentu. Pengujian kelayakan model regresi Poisson dilakukan dengan menggunakan metode *maximum likelihood ratio test* (MLRT) yang memiliki hipotesis sebagai berikut.

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$$

$$H_1 : \text{minimal ada satu } k \text{ sehingga } \beta_k \neq 0 \quad ; k = 1, 2, \dots, p$$

Statistik uji :

$$D(\hat{\beta}) = -2 \ln \Delta = -2 \ln \left(\frac{L(\hat{\omega})}{L(\hat{\Omega})} \right) \quad (2.7)$$

Dimana $D(\hat{\beta})$ adalah nilai devians model regresi Poisson dan $L(\hat{\omega})$ merupakan suatu fungsi *likelihood* untuk model sederhana tanpa melibatkan variabel prediktor, sedangkan $L(\hat{\Omega})$ merupakan suatu fungsi *likelihood* untuk model yang lebih lengkap yang melibatkan variabel prediktor.

Diketahui daerah penolakan adalah Tolak H_0 jika nilai $D(\hat{\beta}) > \chi^2_{(\alpha; p)}$ yang artinya paling sedikit ada satu variabel yang memiliki pengaruh yang signifikan terhadap variabel respon (y). Semakin kecil nilai $D(\hat{\beta})$ maka semakin kecil pula tingkat

kesalahan yang dihasilkan model berarti model yang dihasilkan semakin tepat.

Kemudian dilakukan pengujian parameter secara parsial untuk melihat signifikansi parameter terhadap model dengan hipotesis sebagai berikut.

$$H_0 : \beta_k = 0$$

$$H_1 : \beta_k \neq 0 \quad ; k = 1, 2, \dots, p$$

Statistik uji yang digunakan mengikuti distribusi z yaitu,

$$Z = \frac{\hat{\beta}_k}{SE(\hat{\beta}_k)} \quad (2.8)$$

$(SE(\hat{\beta}_k))^2$ adalah standar error, didapatkan dari elemen diagonal ke- $(k+1)$ dari $\text{var}(\hat{\boldsymbol{\beta}})$ dengan $\text{var}(\hat{\boldsymbol{\beta}}) = -\mathbf{E}(\mathbf{H}^{-1}(\hat{\boldsymbol{\beta}}))$. H_0 akan ditolak jika $|Z|$ lebih besar dari nilai $Z_{\alpha/2}$ dimana α (0,05) adalah tingkat signifikansi yang digunakan.

2.4 Overdispersi Regresi Poisson

Regresi Poisson dikatakan overdispersi apabila nilai variansnya lebih besar dari nilai rata-ratanya. Jika pada data diskrit terjadi overdispersi dan tetap menggunakan regresi Poisson sebagai metode penyelesaiannya tidak efektif karena nilai *standart error* menjadi *under estimate*. Hal ini disebabkan karena parameter koefisien regresi yang dihasilkan dari regresi Poisson tidak efisien meskipun koefisien regresinya tetap konsisten. Overdispersi merupakan nilai dispersi *pearson Chi-square* atau *deviance* dibagi dengan derajat bebasnya yang didapatkan dari banyaknya parameter $L(\hat{\Omega})$ dikurangi banyaknya parameter $L(\hat{\omega})$, diperoleh nilai lebih besar dari 1. Misalkan θ merupakan parameter dispersi, maka jika $\theta > 1$ artinya terjadi overdispersi pada regresi Poisson, jika $\theta < 1$ artinya terjadi underdispersi dan jika $\theta = 1$ berarti tidak terjadi kasus over/underdispersi yang disebut dengan equidispersi (Famoye, Wulu, dan Singh, 2004).

2.5 Regresi Binomial Negatif

Sebagai penerapan dari GLM, maka distribusi binomial negatif memiliki tiga komponen yaitu komponen random, komponen sistematis dan fungsi link (Greene, 2008). Untuk membentuk suatu model regresi diperoleh rata-rata dan varians dalam bentuk:

$$E(Y) = \mu \text{ dan } V[Y] = \mu + \theta\mu^2$$

dengan fungsi massa peluang binomial negatif :

$$f(y, \mu, \theta) = \frac{\Gamma(y + \frac{1}{\theta})}{\Gamma(\frac{1}{\theta})y!} \left(\frac{1}{1 + \theta\mu} \right)^{\frac{1}{\theta}} \left(\frac{\theta\mu}{1 + \theta\mu} \right)^y \quad (2.9)$$

Saat $\theta = 0$ maka distribusi binomial negatif memiliki varians $V[Y] = \mu$. Distribusi binomial negatif akan mendekati suatu distribusi poisson yang mengasumsikan mean dan varians sama yaitu $E[Y] = V[Y] = \mu$. Fungsi distribusi keluarga eksponensial dari distribusi binomial negatif (Greene, 2008) adalah:

$$f(y, \mu, \theta) = \exp \left\{ y \ln \left(\frac{\theta\mu}{1 + \theta\mu} \right) + \frac{1}{\theta} \ln \left(\frac{1}{1 + \theta\mu} \right) + \ln \left(\frac{\Gamma(y + \frac{1}{\theta})}{\Gamma(\frac{1}{\theta})y!} \right) \right\} \quad (2.10)$$

Kontribusi variabel prediktor dalam model regresi binomial negatif dinyatakan dalam bentuk kombinasi linier antara parameter (η) dengan parameter regresi yang akan diestimasi yaitu:

$$\eta_i = \beta_0 + \sum_{k=1}^p \beta_k x_{ik} \quad (2.11)$$

Atau dalam matriks dituliskan dalam bentuk

$$\boldsymbol{\eta} = \mathbf{X}\boldsymbol{\beta} \quad (2.12)$$

dengan $\boldsymbol{\eta}$ adalah vektor ($n \times 1$) dari observasi, \mathbf{X} adalah matriks ($n \times c$) dari variabel prediktor, $\boldsymbol{\beta}$ adalah matriks ($c \times 1$) dari koefisien regresi dengan $c = p + 1$. Nilai ekspektasi variabel respon Y adalah diskrit dan bernilai positif. Maka untuk mentransformasikan nilai η_i (bilangan riil) ke rentang yang sesuai

dengan rentang pada respon Y yang diperlukan suatu fungsi link $g(\cdot)$ (Greene, 2008) yaitu:

$$g(\mu_i) = \ln \mu_i = \mathbf{X}^T \boldsymbol{\beta} \quad (2.13)$$

2.5.1 Estimasi Parameter Regresi Binomial Negatif (RBN)

Estimasi parameter dari regresi binomial negatif digunakan metode maksimum *likelihood* dengan prosedur Newton Raphson. Metode ini membutuhkan turunan pertama dan kedua dari fungsi *likelihood* y_i mempunyai fungsi massa probabilitas distribusi binomial negatif sebagai berikut ini:

$$f(y_i | \mu_i, \theta) = \frac{\Gamma(y_i + 1/\theta)}{\Gamma(1/\theta)\Gamma(y_i + 1)} \left(\frac{1}{1 + \theta\mu_i} \right)^{1/\theta} \left(\frac{\theta\mu_i}{1 + \theta\mu_i} \right)^{y_i} \quad i = 0, 1, 2, \dots \quad (2.14)$$

Karena fungsinya saling bebas, maka fungsi *log-likelihood* adalah:

$$L(\boldsymbol{\beta}, \theta) = \prod_{i=1}^n \frac{\Gamma(y_i + 1/\theta)}{\Gamma(1/\theta)\Gamma(y_i + 1)} \left(\frac{1}{1 + \theta\mu_i} \right)^{1/\theta} \left(\frac{\theta\mu_i}{1 + \theta\mu_i} \right)^{y_i}$$

Dengan $\frac{\Gamma(y_i + 1/\theta)}{\Gamma(1/\theta)} = \prod_{r=1}^n (r + \theta^{-1})$

$$L(\boldsymbol{\beta}, \theta) = \prod_{i=1}^n \left(\prod_{r=1}^{y_i-1} (r + \theta^{-1}) \right) \frac{1}{(y_i!)} \left(\frac{1}{1 + \theta\mu_i} \right)^{1/\theta} \left(\frac{\theta\mu_i}{1 + \theta\mu_i} \right)^{y_i}$$

$$\ln\{L(\boldsymbol{\beta}, \theta)\} = \sum_{i=1}^n [\Lambda]$$

dimana

$$\Lambda = \left(\sum_{r=1}^{y_i-1} \ln(r + \theta^{-1}) \right) - \ln(y_i!) + y_i \ln(\theta\mu_i) - (\theta^{-1} + y_i) \ln(1 + \theta\mu_i)$$

Turunan pertama dari fungsi *log-likelihood* terhadap koefisien $\boldsymbol{\beta}$ adalah:

$$\begin{aligned}
\frac{\partial L(\boldsymbol{\beta}, \theta)}{\partial \beta_0} &= \sum_{i=1}^n \left[(y_i + \theta^{-1}) \left(\frac{\theta \mu_i}{1 + \theta \mu_i} \right) \right] = \sum_{i=1}^n \left[\frac{y_i - \mu_i}{1 + \theta \mu_i} \right] = 0 \\
&\vdots \\
\frac{\partial L(\boldsymbol{\beta}, \theta)}{\partial \beta_p} &= \sum_{i=1}^n \left[y_i x_{ip} - (y_i + \theta^{-1}) \left(\frac{\theta \mu_i x_{ip}}{1 + \theta \mu_i} \right) \right] \\
&= \sum_{i=1}^n \left[\frac{(y_i - \mu_i) x_{ip}}{1 + \theta \mu_i} \right] \\
&= \sum_{i=1}^n \left[\frac{\mu_i}{1 + \theta \mu_i} \frac{(y_i - \mu_i) x_{ip}}{\mu_i} \right] = 0
\end{aligned}$$

Bentuk persamaan matriks dari turunan pertama fungsi *log-likelihood* terhadap parameter β yaitu: $\mathbf{q} = \mathbf{X}^T \mathbf{W} \mathbf{z}$, dengan \mathbf{X} adalah matriks $(n \times c)$ dari variabel prediktor, \mathbf{W} adalah matriks pembobot diagonal ke- i dan \mathbf{z} adalah vektor matriks dengan baris ke- i , dengan masing masing elemennya adalah:

$$w_i = \frac{\mu_i}{1 + \theta \mu_i} \quad \text{dan} \quad z_i = \frac{(y_i - \mu_i)}{\mu_i} \quad i = 1, 2, \dots, n$$

Turunan pertama dari fungsi *log-likelihood* terhadap parameter dispersi θ adalah:

$$\begin{aligned}
f'(\theta) = \frac{\partial L(\boldsymbol{\beta}, \theta)}{\partial \theta} &= \sum_{i=1}^n \left[-\theta^{-2} \sum_{r=0}^{y_i-1} \frac{1}{r + \theta^{-1}} + \frac{y_i}{\theta} + \theta^{-2} \ln(1 + \theta \mu_i) \right. \\
&\quad \left. - \frac{(\theta^{-1} + y_i) \mu_i}{1 + \theta \mu_i} \right] = 0 \\
f'(\theta) = \frac{\partial L(\boldsymbol{\beta}, \theta)}{\partial \theta} &= \sum_{i=1}^n \left[-\theta^{-2} \sum_{r=0}^{y_i-1} \frac{1}{r + \theta^{-1}} + \theta^{-2} \ln(1 + \theta \mu_i) + \frac{y_i - \mu_i}{\theta(1 + \theta \mu_i)} \right] = 0
\end{aligned}$$

Turunan parsial kedua fungsi *log-likelihood* terhadap parameter koefisien regresi β adalah:

$$\frac{\partial^2 L(\boldsymbol{\beta}, \theta)}{\partial \beta_0^2} = - \sum_{i=1}^n \left[\frac{(1 + \theta y_i \mu_i)}{(1 + \theta \mu_i)^2} \right]$$

$$\frac{\partial^2 L(\boldsymbol{\beta}, \theta)}{\partial \beta_0 \partial \beta_k} = \sum_{i=1}^n \left[\frac{-\mu_i x_{ik} (1 + \theta \mu_i) - (y_i - \mu_i)(x_{ik} \theta \mu_i)}{(1 + \theta \mu_i)^2} \right]$$

$$\frac{\partial^2 L(\boldsymbol{\beta}, \theta)}{\partial \beta_0 \partial \beta_k} = \sum_{i=1}^n \left[\frac{(1 + \theta y_i) x_{ik} \mu_i}{(1 + \theta \mu_i)^2} \right]$$

Misalkan turunan parsial pertama dari $L(\boldsymbol{\beta}, \theta)$ terhadap $\beta_k, k \leq p$ adalah:

$$\frac{\partial^2 L(\boldsymbol{\beta}, \theta)}{\partial \beta_k} = - \sum_{i=1}^n \left[\frac{(y_i - \mu_i) x_{ik}}{(1 + \theta \mu_i)} \right] = 0, \text{ maka turunan parsial kedua}$$

terhadap $\beta_u, u \leq p$ adalah:

$$\frac{\partial^2 L(\boldsymbol{\beta}, \theta)}{\partial \beta_u \partial \beta_k} = \sum_{i=1}^n \left[\frac{-\mu_i x_{iu} x_{ik} (1 + \theta \mu_i) - (y_i - \mu_i)(x_{iu} x_{ik} \theta \mu_i)}{(1 + \theta \mu_i)^2} \right]$$

$$\frac{\partial^2 L(\boldsymbol{\beta}, \theta)}{\partial \beta_0 \partial \beta_k} = - \sum_{i=1}^n \left[\frac{(1 + \theta y_i) x_{iu} x_{ik} \mu_i}{(1 + \theta \mu_i)^2} \right]$$

Ekspektasi dari turunan kedua *log-likelihood* adalah:

$$E \left(- \frac{\partial^2 L(\boldsymbol{\beta}, \theta)}{\partial \beta_u \partial \beta_k} \right) = \sum_{i=1}^n \left[\frac{x_{iu} x_{ik} \mu_i}{(1 + \theta \mu_i)} \right] \quad (2.15)$$

Jika persamaan (2.15) dinyatakan dalam matriks \mathbf{I} yaitu matriks yang mengandung ekspektasi negatif dari turunan kedua *log-likelihood* maka:

$$\mathbf{I} = \mathbf{X}^T \mathbf{W} \mathbf{X}$$

dengan \mathbf{X} adalah matriks dari variabel prediktor, \mathbf{W} adalah matriks pembobot diagonal ke- i dengan elemen : $w_i = \frac{\mu_i}{1 + \theta \mu_i}$

Turunan kedua fungsi *log-likelihood* terhadap parameter dispersi θ adalah:

$$f''(\theta) = \frac{\partial^2 L(\boldsymbol{\beta}, \theta)}{\partial \theta^2} = \sum_{i=1}^n \left[\theta^{-3} \sum_{r=0}^{y_i-1} \frac{(2r + \theta^{-1})}{(r + \theta^2 \mu_i)^2} - 2\theta^{-3} \ln(1 + \theta \mu_i) \right. \\ \left. + \frac{-\theta^2 \mu_i}{(1 + \theta \mu_i)} + \frac{(y_i - \mu_i)(1 + 2\theta \mu_i)}{(\theta + \theta^2 \mu_i)^2} \right]$$

Langkah langkah estimasi parameter regresi binomial negatif dilakukan dengan langkah sebagai berikut:

1. Menentukan estimasi awal θ , misal $\hat{\theta}_1 = 0,1$
2. Menentukan estimasi maksimum *likelihood* dari parameter $\boldsymbol{\beta}$ menggunakan prosedur iterasi Fisher scoring dengan asumsi $\theta = \hat{\theta}_1$

$$\hat{\boldsymbol{\beta}}_{i+1} = \hat{\boldsymbol{\beta}}_i + (\mathbf{X}^T \mathbf{W}_i \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}_i \mathbf{z}_i$$

Iterasi berakhir jika diperoleh $\|\hat{\boldsymbol{\beta}}_{i+1} - \hat{\boldsymbol{\beta}}_i\| \leq \varepsilon$

3. Menggunakan $\hat{\boldsymbol{\beta}}$ untuk menghasilkan estimasi dari parameter θ dengan menggunakan prosedur iterasi Newton-Raphson satu variabel, iterasi berakhir jika diperoleh $\|\hat{\theta}_{i+1} - \hat{\theta}_i\| \leq \varepsilon$

$$\hat{\theta}_{i+1} = \hat{\theta}_i - \frac{f'(\theta_i)}{f''(\theta_i)}$$

4. Jika $|\hat{\theta}_{i+1} - \hat{\theta}_i| \leq \varepsilon$ selesai; bila tidak, gunakan parameter $\theta = \hat{\theta}_{i+1}$ dan kembali ke langkah 2, nilai ε merupakan bilangan positif yang sangat kecil.

2.5.2 Pengujian Parameter Regresi Binomial Negatif

Pengujian signifikansi secara serentak untuk estimasi parameter model regresi Binomial Negatif menggunakan uji devians dengan hipotesis sebagai berikut (Hosmer, Lemeshow, dan Sturdivant, 2013).

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$$

$$H_1 : \text{minimal ada satu } \beta_k \neq 0 \quad ; k = 1, 2, \dots, p$$

Statistik Uji:

$$D(\hat{\beta}) = -2 \ln \left(\frac{L(\hat{\omega})}{L(\hat{\Omega})} \right) = 2 (\ln L(\hat{\Omega}) - \ln L(\hat{\omega})) \quad (2.16)$$

Tolak H_0 jika statistik uji $D(\hat{\beta}) > \chi^2_{(\alpha; p)}$

Pengujian signifikansi secara parsial untuk mengetahui parameter mana saja yang memberikan pengaruh yang signifikan terhadap model dengan hipotesis sebagai berikut.

$$H_0 : \beta_k = 0$$

$$H_1 : \beta_k \neq 0; k = 1, 2, \dots, p$$

Statistik Uji:

$$z_{hit} = \frac{\hat{\beta}_k}{SE(\hat{\beta}_k)} \quad (2.17)$$

Tolak H_0 jika statistik uji $|z_{hit}|$ lebih besar dari nilai $Z_{\alpha/2}$, yang artinya bahwa parameter ke- k signifikan terhadap model regresi Binomial Negatif.

2.6 Pengujian Aspek Data Spasial

Regresi spasial merupakan salah satu metode yang digunakan untuk mengetahui hubungan antara variabel respon dan variabel prediktor dengan memperhatikan aspek lokasi atau spasial. Aspek spasial yang dimaksud adalah data yang digunakan memiliki *error* saling berkorelasi dan memiliki heterogenitas spasial (Anselin, 1988).

2.6.1 Pengujian Dependensi Spasial

Pengujian dependensi spasial dilakukan untuk melihat apakah pengamatan disuatu lokasi berpengaruh terhadap pengamatan dilokasi lain yang letaknya berdekatan. Pengujian dependensi spasial dilakukan menggunakan statistik uji Moran's I dengan hipotesis sebagai berikut (Anselin, 1988).

$$H_0 : I = 0 \quad (\text{tidak terdapat dependensi spasial})$$

$$H_1 : I \neq 0 \quad (\text{terdapat dependensi spasial})$$

Statistik uji :

$$Z_{hit} = \frac{\hat{I} - E(\hat{I})}{\sqrt{Var(\hat{I})}} \quad (2.18)$$

dengan

$$\hat{I} = \frac{n \sum_{i=1}^n \sum_{j=1}^n w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\left(\sum_{i=1}^n \sum_{j=1}^n w_{ij} \right) \sum_{i=1}^n (y_i - \bar{y})^2} \quad (2.19)$$

$$E(\hat{I}) = -\frac{1}{n-1}; Var(\hat{I}) = \frac{n^2 s - n s_2 + 3 s_0^2}{(n^2 - 1) s_0^2} - E(\hat{I})^2$$

$$s_0^2 = \sum_{i=1}^n \sum_{j=1}^n w_{ij}$$

$$s_1 = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (w_{ij} + w_{ji})^2$$

$$s_2 = \sum_{i=1}^n \left(\sum_{j=1}^n w_{ij} + \sum_{j=1}^n w_{ji} \right)^2$$

dimana,

n = banyak pengamatan

\bar{y} = nilai rata-rata dari y_i dari n lokasi

y_i = nilai pengamatan pada lokasi ke-i

y_j = nilai pengamatan pada lokasi ke-j

w_{ij} = elemen matriks pembobot spasial

Keputusan : Tolak H_0 jika nilai $|Z_{hit}|$ lebih besar dari nilai $Z_{\alpha/2}$, yang berarti terdapat dependensi spasial dalam model.

2.6.2 Pengujian Heterogenitas Spasial

Menurut Anselin (1988) pengujian heterogenitas spasial dilakukan untuk melihat apakah terdapat kekhasan pada setiap lokasi pengamatan, sehingga parameter regresi yang dihasilkan berbeda-beda secara spasial. Pengujian heterogenitas spasial dilakukan menggunakan statistik uji *Breusch-Pagan* (BP) dengan hipotesis sebagai berikut.

$$H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2 \text{ (varians antarlokasi sama)}$$

$$H_1: \text{minimal ada satu } \sigma_i^2 \neq \sigma^2 \text{ (varians antarlokasi berbeda)}$$

Statistik uji:

$$BP = \left(\frac{1}{2} \right) \mathbf{f}^T \mathbf{Z} (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{f} \sim \chi^2_{(p)} \quad (2.20)$$

dimana

$$e_i = y_i - \hat{y}_i$$

$$\mathbf{f} = (f_1, f_2, \dots, f_n)^T \text{ dengan } f_i = \frac{e_i^2}{\hat{\sigma}^2} - 1$$

$$e_i^2 = \text{kuadrat sisaan untuk pengamatan ke-} i$$

\mathbf{Z} = matriks berukuran $n \times (p+1)$ berisi vektor yang sudah dinormalisasikan untuk setiap pengamatan

$$\hat{\sigma}^2 = \text{varians } y$$

Keputusan : Tolak H_0 jika nilai $BP > \chi^2_{(\alpha, k)}$ atau $p\text{-value} < \alpha$ (0,05) yang berarti varians antar lokasi berbeda.

2.7 Model Geographically Weighted Negative Binomial Regression (GWNBR)

Model GWNBR akan menghasilkan parameter lokal dengan masing-masing lokasi akan memiliki parameter yang berbeda-beda. Model GWNBR dapat dirumuskan sebagai berikut (Ricardo dan Carvalho, 2013).

$$y_i \sim NB \left[t_j \exp \left(\beta_0 + \sum_{k=1}^p \beta_k(u_i, v_i) x_{ik} \right), \theta(u_i, v_i) \right], i=1,2,3,\dots,n \quad (2.21)$$

dimana,

y_i : nilai observasi respon ke- i

t_j : *offset variable*

x_{ik} : nilai observasi variabel prediktor ke- k pada pengamatan lokasi (u_i, v_i)

$\beta_k(u_i, v_i)$: koefisien regresi variabel prediktor ke- k untuk setiap lokasi (u_i, v_i)

$\theta(u_i, v_i)$: parameter dispersi untuk setiap lokasi (u_i, v_i)

Fungsi sebaran binomial negatif untuk setiap lokasi berdasarkan persamaan (2.21) dapat ditulis dalam bentuk persamaan berikut:

$$f(y_i | x_{ik} \boldsymbol{\beta}_{(u_i, v_i)}, \theta_{(u_i, v_i)}) = \frac{\Gamma y_i + 1/\theta_i}{\Gamma(1/\theta_i) \Gamma(y_i + 1)} \left(\frac{1}{1 + \theta_i + \mu_i} \right)^{1/\theta_i} \left(\frac{\theta_i \mu_i}{1 + \theta_i \mu_i} \right)^{y_i} \quad (2.22)$$

$i = 1, 2, \dots, n$

dimana:

$$\mu_i = \exp(\mathbf{x}_i^T \boldsymbol{\beta}_{(u_i, v_i)})$$

$$\theta_i = \theta_{(u_i, v_i)}$$

Penaksir parameter koefisien GWNBR dilakukan dengan metode maksimum *likelihood*. Faktor letak geografis merupakan faktor pembobot pada model GWNBR yang memiliki nilai yang berbeda beda untuk setiap lokasi dan menunjukkan sifat lokal pada model. Fungsi *likelihood* yang telah diberi pembobot adalah sebagai berikut (Richardo & Carvalho, 2013) :

$$\ln L \boldsymbol{\beta}_{(u_i, v_i)}, \theta_{(u_i, v_i)} | y_i, \mathbf{x}_{ik} = \sum_{j=1}^n W_{j(u_i, v_i)}(\Delta) \quad (2.23)$$

dimana

$$\Delta = \ln \frac{\Gamma(y_i + 1/\theta_i)}{\Gamma(1/\theta_i) \Gamma(y_i + 1)} + y_i \ln(\theta_i \mu_i) - (\theta^{-1} + y_i) \ln(1 + \theta_i \mu_i)$$

Proses estimasi parameter koefisien regresi diperoleh melalui metode iterasi numerik Newton Raphson.

2.7.1 Estimasi Parameter Model GWNBR

Model GWNBR merupakan pengembangan dari model regresi binomial negatif. Model ini menghasilkan estimasi parameter model yang bersifat lokal untuk setiap titik atau lokasi di mana data tersebut dikumpulkan. Dalam model GWNBR variabel respon yang diprediksi dengan variabel prediktor yang masing-masing koefisien regresinya bergantung pada lokasi di mana data tersebut diamati. Model GWNBR yang merupakan pengembangan dari distribusi binomial negatif (2.9) yaitu:

$$f(y_i | x_{ik} \beta_{(u_i, v_i)}, \theta_i) = \frac{\Gamma y_i + 1/\theta_i}{\Gamma(1/\theta_i) \Gamma(y_i + 1)} \left(\frac{1}{1 + \theta_i \mu_i} \right)^{1/\theta_i} \left(\frac{\theta_i \mu_i}{1 + \theta_i \mu_i} \right)^{y_i} \quad (2.24)$$

dengan $i = 1, 2, \dots, n$

dimana:

$$\mu_i = \exp(\mathbf{x}_i^T \boldsymbol{\beta}_{(u_i, v_i)})$$

$$\theta_i = \theta_{(u_i, v_i)}$$

Estimasi parameter model GWNBR dilakukan dengan menggunakan metode *Maximum Likelihood Estimation* (MLE). Langkah awal dari metode ini adalah membentuk fungsi *likelihood* yaitu sebagai berikut:

Fungsi *likelihood*:

$$\begin{aligned} L(\boldsymbol{\beta}_{(u_i, v_i)}, \theta_i | y_i, x_i) &= \prod_{i=1}^n \left[\frac{\Gamma(y_i + 1/\theta_i)}{\Gamma(1/\theta_i) \Gamma(y_i + 1)} \left(\frac{1}{1 + \theta_i \mu_i} \right)^{1/\theta_i} \left(\frac{\theta_i \mu_i}{1 + \theta_i \mu_i} \right)^{y_i} \right] \\ &= \prod_{i=1}^n \frac{\Gamma(y_i + 1/\theta_i)}{\Gamma(1/\theta_i) \Gamma(y_i + 1)} \left(\prod_{i=1}^n \left(\frac{1}{1 + \theta_i \mu_i} \right)^{1/\theta_i} \right) \left(\prod_{i=1}^n \left(\frac{\theta_i \mu_i}{1 + \theta_i \mu_i} \right)^{y_i} \right) \end{aligned}$$

Diketahui bahwa $\frac{\Gamma(y+c)}{\Gamma(c)} = c(c+1)(c+2)\dots(c+y-1)$,

sehingga diperoleh:

$$\begin{aligned}\frac{\Gamma(y_i + \theta_i^{-1})}{\Gamma(\theta_i^{-1})} &= \left(\frac{1}{\theta_i}\right) \left(\frac{1}{\theta_i} + 1\right) \left(\frac{1}{\theta_i} + 2\right) \dots \left(\frac{1}{\theta_i} + y_i - 1\right) \\ &= \left(\frac{1}{\theta_i}\right) \left(\frac{1}{\theta_i}\right) (1 + \theta_i) \left(\frac{1}{\theta_i}\right) (1 + 2\theta_i) \dots \left(\frac{1}{\theta_i}\right) (1 + (y_i - 1)\theta_i) \\ &= \Pi_{r=0}^{y_i-1} (r + \theta_i^{-1})\end{aligned}$$

Maka fungsi *likelihood* $L(\boldsymbol{\beta}_{(u_i, v_i)}, \theta_i | y_i, x_i)$ jika disederhanakan adalah sebagai berikut:

$$L(\boldsymbol{\beta}_{(u_i, v_i)}, \theta_i | y_i, x_i) = \Pi_{i=1}^n \left(\Pi_{r=0}^{y_i-1} (r + \theta_i^{-1}) \right) \frac{1}{y_i!} \left(\frac{1}{1 + \theta_i \mu_i} \right)^{1/\theta_i} \left(\frac{\theta_i \mu_i}{1 + \theta_i \mu_i} \right)^{y_i}$$

Kemudian fungsi *likelihood* tersebut diubah dalam bentuk logaritma natural menjadi:

$$\begin{aligned}\ln L_{\boldsymbol{\beta}_{(u_i, v_i)}, \theta_i | y_i, x_i} &= \\ &= \sum_{i=1}^n \left[\left(\sum_{r=0}^{y_i-1} \ln(r + \theta_i^{-1}) \right) - \ln(y_i!) + \theta_i^{-1} \ln(1 + \theta_i \mu_i) + y_i \ln \left(\frac{\theta_i \mu_i}{1 + \theta_i \mu_i} \right) \right] \\ &= \sum_{i=1}^n \left[\left(\sum_{r=0}^{y_i-1} \ln(r + \theta_i^{-1}) \right) - \ln(y_i!) + y_i \ln \theta_i \mu_i - (y_i + \theta_i^{-1}) \ln(1 + \theta_i \mu_i) \right]\end{aligned}$$

dengan $\mu_i = \exp(\mathbf{x}_i^T \boldsymbol{\beta}_{(u_i, v_i)})$ maka diperoleh bentuk lain dari persamaan logaritma fungsi *likelihood* menjadi:

$$L(\boldsymbol{\beta}_{(u_i, v_i)}, \theta_i | y_i, x_i) = \sum_{i=1}^n [a + b - c]$$

Faktor letak geografis merupakan faktor pembobot pada model GWNBR. Faktor ini memiliki nilai yang berbeda untuk setiap wilayah yang menunjukkan sifat lokasi pada model GWNBR. Oleh karena itu, pembobot diberikan pada bentuk persamaan *ln-likelihood* untuk model GWNBR, sehingga diperoleh bentuk persamaan dibawah ini:

$$L(\boldsymbol{\beta}_{(u_i, v_i)}, \theta_i | y_i, x_i) = \sum_{i=1}^n w_{j_{(u_i, v_i)}} [a + b - c]$$

Dengan keterangan:

$$\begin{aligned}
a &= \left(\sum_{r=0}^{y_j-1} \ln(r + \theta_i^{-1}) \right) - \ln(y_i!) \\
b &= y_j \ln(\theta_i \exp(\mathbf{x}_i^T \boldsymbol{\beta}_{(u_i, v_i)})) \\
c &= (y_j + \theta_j^{-1}) \ln(1 + \theta_j \exp(\mathbf{x}_j^T \boldsymbol{\beta}_{(u_i, v_i)}))
\end{aligned}$$

Turunan pertama dari logaritma fungsi *likelihood* terhadap parameter dispersi θ adalah:

$$\begin{aligned}
\frac{\partial \ln L(.)}{\partial \theta_i} &= \sum_{i=1}^n w_j(u_i, v_i) \left[\frac{\partial a}{\partial \theta_i} + \frac{\partial b}{\partial \theta_i} - \frac{\partial c}{\partial \theta_i} \right] \\
\frac{\partial a}{\partial \theta_i} &= -\theta_i^{-2} \sum_{r=0}^{y_i-1} \frac{1}{(r + \theta_i^{-1})} \\
\frac{\partial b}{\partial \theta_i} &= y_i \frac{1}{\theta_i \exp(\mathbf{x}_i^T \boldsymbol{\beta}_{(u_i, v_i)})} \left(\exp(\mathbf{x}_i^T \boldsymbol{\beta}_{(u_i, v_i)}) \right) = \frac{y_i}{\theta_i} \\
\frac{\partial c}{\partial \theta_i} &= \left(-\frac{\ln(1 + \theta_i \exp(\mathbf{x}_i^T \boldsymbol{\beta}_{(u_i, v_i)}))}{\theta_i^2} + \frac{(y_i + \theta_i^{-1}) \exp(\mathbf{x}_i^T \boldsymbol{\beta}_{(u_i, v_i)})}{1 + \theta_i \exp(\mathbf{x}_i^T \boldsymbol{\beta}_{(u_i, v_i)})} \right) \\
&= (-d + e)
\end{aligned}$$

Dimana

$$d = \frac{\ln(1 + \theta_i \exp(\mathbf{x}_i^T \boldsymbol{\beta}_{(u_i, v_i)}))}{\theta_i^2} \text{ dan } e = \frac{(y_i + \theta_i^{-1}) \exp(\mathbf{x}_i^T \boldsymbol{\beta}_{(u_i, v_i)})}{1 + \theta_i \exp(\mathbf{x}_i^T \boldsymbol{\beta}_{(u_i, v_i)})}$$

Sehingga

$$\frac{\partial \ln L(.)}{\partial \theta_i} = \sum_{i=1}^n w_i(u_i, v_i) \left[\frac{\partial a}{\partial \theta_i} + \frac{\partial b}{\partial \theta_i} + d - e \right]$$

Turunan kedua dari logaritma fungsi *likelihood* terhadap parameter dispersi θ adalah :

$$\begin{aligned}
\frac{\partial^2 \ln L(.)}{\partial \theta_i^2} &= \sum_{i=1}^n w_i(u_i, v_i) \left[\frac{\partial^2 a}{\partial \theta_i^2} + \frac{\partial^2 b}{\partial \theta_i^2} - \frac{\partial^2 c}{\partial \theta_i^2} \right] \\
&= \sum_{i=1}^n w_i(u_i, v_i) \left[\frac{\partial^2 a}{\partial \theta_i^2} + \frac{\partial^2 b}{\partial \theta_i^2} - \left(-\frac{\partial d}{\partial \theta_i} + \frac{\partial e}{\partial \theta_i} \right) \right]
\end{aligned}$$

$$\frac{\partial^2 \ln L(.)}{\partial \theta_i^2} = \sum_{i=1}^n w_{i(u_i, v_i)} \left[\frac{\partial^2 a}{\partial \theta_i^2} + \frac{\partial^2 b}{\partial \theta_i^2} + \frac{\partial d}{\partial \theta_i} + \frac{\partial e}{\partial \theta_i} \right]$$

Dengan keterangan:

$$\begin{aligned} \frac{\partial^2 a}{\partial \theta_i^2} &= \sum_{r=0}^{y_i-1} \frac{2\theta_i^{-3} (r + \theta_i^{-1}) - \theta_i^{-4}}{(r + \theta_i^{-1})^2} \\ &= \sum_{r=0}^{y_i-1} \frac{2\theta_i^{-3} r + 2\theta_i^{-4} - \theta_i^{-4}}{(r + \theta_i^{-1})^2} \\ &= \sum_{r=0}^{y_i-1} \theta_i^{-3} \frac{(2r + \theta_i^{-1})}{(r + \theta_i^{-1})^2} \end{aligned}$$

$$\frac{\partial^2 b}{\partial \theta_i^2} = -\frac{y_i}{\theta_i^2}$$

$$\begin{aligned} \frac{\partial d}{\partial \theta_i} &= \left(\frac{\theta_i^2 \frac{\partial}{\partial \theta_i} \ln(\theta_i \exp(\mathbf{x}_i^T \boldsymbol{\beta}_{(u_i, v_i)})) - \ln(\theta_i \exp(\mathbf{x}_i^T \boldsymbol{\beta}_{(u_i, v_i)})) \frac{\partial}{\partial \theta_i} \theta_i^2}{\theta_i^4} \right) \\ &= \frac{\exp(\mathbf{x}_i^T \boldsymbol{\beta}_{(u_i, v_i)})}{(1 + \theta_i \exp(\mathbf{x}_i^T \boldsymbol{\beta}_{(u_i, v_i)})) \theta_i^2} - \frac{2 \ln(1 + \theta_i) \exp(\mathbf{x}_i^T \boldsymbol{\beta}_{(u_i, v_i)})}{\theta_i^3} \\ &\quad - \frac{\exp(\mathbf{x}_i^T \boldsymbol{\beta}_{(u_i, v_i)})}{\theta_i^2} (1 + \theta_i \exp(\mathbf{x}_i^T \boldsymbol{\beta}_{(u_i, v_i)})) - (\exp(\mathbf{x}_i^T \boldsymbol{\beta}_{(u_i, v_i)}))^2 (y_i + \theta_i^{-1}) \\ \frac{\partial e}{\partial \theta_i} &= \frac{(\exp(\mathbf{x}_i^T \boldsymbol{\beta}_{(u_i, v_i)}))^2 (y_i + \theta_i^{-1})}{(1 + \theta_i \exp(\mathbf{x}_i^T \boldsymbol{\beta}_{(u_i, v_i)}))^2} \\ &= -\frac{\exp(\mathbf{x}_i^T \boldsymbol{\beta}_{(u_i, v_i)})}{(1 + \theta_i \exp(\mathbf{x}_i^T \boldsymbol{\beta}_{(u_i, v_i)}))^2 \theta_i^2} - \frac{(\exp(\mathbf{x}_i^T \boldsymbol{\beta}_{(u_i, v_i)}))^2 (y_i + \theta_i^{-1})}{(1 + \theta_i \exp(\mathbf{x}_i^T \boldsymbol{\beta}_{(u_i, v_i)}))^2 \theta_i^2} \end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 c}{\partial \theta_i^2} &= \frac{\partial d}{\partial \theta_i} - \frac{\partial e}{\partial \theta_i} \\
&= 2 \left(\frac{\exp(\mathbf{x}_i^T \boldsymbol{\beta}_{(u_i, v_i)})}{(1 + \theta_i \exp(\mathbf{x}_i^T \boldsymbol{\beta}_{(u_i, v_i)}))^2 \theta_i^2} \right) - \frac{2 \ln(1 + \theta_i) \exp(\mathbf{x}_i^T \boldsymbol{\beta}_{(u_i, v_i)})}{\theta_i^3} \\
&\quad + \frac{(\exp(\mathbf{x}_i^T \boldsymbol{\beta}_{(u_i, v_i)}))^2 (y_i + \theta_i^{-1})}{(1 + \theta_i \exp(\mathbf{x}_i^T \boldsymbol{\beta}_{(u_i, v_i)}))^2 \theta_i^2}
\end{aligned}$$

Untuk mendapatkan $\hat{\boldsymbol{\beta}}_{(u_i, v_i)}$ yang merupakan estimasi parameter $\boldsymbol{\beta}_{(u_i, v_i)}$ maka bentuk persamaan $\ln L(\cdot)$ diturunkan terhadap $\boldsymbol{\beta}^T_{(u_i, v_i)}$ adalah:

$$\begin{aligned}
\frac{\partial \ln L(\cdot)}{\partial \boldsymbol{\beta}^T_{(u_i, v_i)}} &= \sum_{i=1}^n w_{i(u_i, v_i)} \left[\frac{\partial a}{\partial \boldsymbol{\beta}^T_{(u_i, v_i)}} + \frac{\partial b}{\partial \boldsymbol{\beta}^T_{(u_i, v_i)}} - \frac{\partial c}{\partial \boldsymbol{\beta}^T_{(u_i, v_i)}} \right] \\
\frac{\partial a}{\partial \boldsymbol{\beta}^T_{(u_i, v_i)}} &= 0 \\
\frac{\partial b}{\partial \boldsymbol{\beta}^T_{(u_i, v_i)}} &= \frac{y_i}{\exp(\boldsymbol{\beta}^T_{(u_i, v_i)} \mathbf{x}_i)} \theta_i \exp(\boldsymbol{\beta}^T_{(u_i, v_i)} \mathbf{x}_i) \mathbf{x}_i \\
\frac{\partial c}{\partial \boldsymbol{\beta}^T_{(u_i, v_i)}} &= (y_i + \theta_i^{-1}) \frac{\theta_i \exp(\boldsymbol{\beta}^T_{(u_i, v_i)} \mathbf{x}_i) \mathbf{x}_i}{1 + \theta_i \exp(\boldsymbol{\beta}^T_{(u_i, v_i)} \mathbf{x}_i)}
\end{aligned}$$

Sehingga, dengan mensubstitusikan nilai-nilai diatas maka dapat disederhanakan menjadi sebagai berikut:

$$\begin{aligned}
\frac{\partial \ln L(\cdot)}{\partial \boldsymbol{\beta}^T_{(u_i, v_i)}} &= \sum_{i=1}^n w_{i(u_i, v_i)} \left[y_i \mathbf{x}_i - (y_i + \theta_i^{-1}) \frac{\theta_i \exp(\mathbf{x}_i \boldsymbol{\beta}^T_{(u_i, v_i)}) \mathbf{x}_i}{1 + \theta_i \exp(\mathbf{x}_i \boldsymbol{\beta}^T_{(u_i, v_i)})} \right] \\
&= \sum_{i=1}^n w_{i(u_i, v_i)} \left[\frac{(y_i - \exp(\mathbf{x}_i \boldsymbol{\beta}^T_{(u_i, v_i)})) \mathbf{x}_i}{1 + \theta_i \exp(\mathbf{x}_i \boldsymbol{\beta}^T_{(u_i, v_i)})} \right]
\end{aligned}$$

Turunan parsial kedua logaritma fungsi *likelihood* terhadap parameter regresi $\boldsymbol{\beta}_{(u_i, v_i)}$ adalah:

$$\frac{\partial^2 \ln L(.)}{\partial \boldsymbol{\beta}^T(u_i, v_i) \partial \boldsymbol{\beta}(u_i, v_i)} = \frac{\partial b}{\partial \boldsymbol{\beta}^T(u_i, v_i) \partial \boldsymbol{\beta}(u_i, v_i)} - \frac{\partial c}{\partial \boldsymbol{\beta}^T(u_i, v_i) \partial \boldsymbol{\beta}(u_i, v_i)}$$

$$\frac{\partial b}{\partial \boldsymbol{\beta}^T(u_i, v_i) \partial \boldsymbol{\beta}(u_i, v_i)} = 0$$

$$\frac{\partial c}{\partial \boldsymbol{\beta}^T(u_i, v_i) \partial \boldsymbol{\beta}(u_i, v_i)} = \left[- \left(y_i + \theta_i^{-1} \right) \theta \mathbf{x}_i \mathbf{x}_i^T \left(\exp \left(\mathbf{x}_i^T \boldsymbol{\beta}(u_i, v_i) \right) \right) (u) \right]$$

Dimana

$$u = \frac{1 + \theta_i \exp \left(\mathbf{x}_i^T \boldsymbol{\beta}(u_i, v_i) \right) - \theta_i \exp \left(\mathbf{x}_i^T \boldsymbol{\beta}(u_i, v_i) \right)}{\left(1 + \theta_i \exp \left(\mathbf{x}_i^T \boldsymbol{\beta}(u_i, v_i) \right) \right)^2} = \frac{1}{\left(1 + \theta_i \exp \left(\mathbf{x}_i^T \boldsymbol{\beta}(u_i, v_i) \right) \right)^2}$$

Sehingga $\frac{\partial^2 \ln L(.)}{\partial \boldsymbol{\beta}^T(u_i, v_i) \partial \boldsymbol{\beta}(u_i, v_i)}$ setelah diberi pembobot menjadi:

$$= \sum_{i=1}^n w_{i(u_i, v_i)} \left[- \left(y_i + \theta_i^{-1} \right) \theta \mathbf{x}_i \mathbf{x}_i^T \left(\exp \left(\mathbf{x}_i^T \boldsymbol{\beta}(u_i, v_i) \right) \right) \frac{1}{\left(1 + \theta_i \exp \left(\mathbf{x}_i^T \boldsymbol{\beta}(u_i, v_i) \right) \right)^2} \right]$$

$$= \sum_{i=1}^n w_{i(u_i, v_i)} \left[\frac{- \left(y_i + \theta_i^{-1} \right) \left(\exp \left(\mathbf{x}_i^T \boldsymbol{\beta}(u_i, v_i) \right) \right) \theta \mathbf{x}_i \mathbf{x}_i^T}{\left(1 + \theta_i \exp \left(\mathbf{x}_i^T \boldsymbol{\beta}(u_i, v_i) \right) \right)^2} \right]$$

Turunan parsial kedua logaritma fungsi *likelihood* terhadap parameter regresi $\boldsymbol{\beta}(u_i, v_i)$ dan parameter dispersi θ_i adalah:

$$\frac{\partial^2 \ln L(.)}{\partial \boldsymbol{\beta}(u_i, v_i) \partial \theta_i} = \sum_{i=1}^n w_{i(u_i, v_i)} \left[\frac{\theta_i^3 - \exp \left(\boldsymbol{\beta}^T(u_i, v_i) \mathbf{x}_i \right) \mathbf{x}_i}{1 + \theta_i \exp \left(\boldsymbol{\beta}^T(u_i, v_i) \mathbf{x}_i \right)} - \frac{\exp \left(\boldsymbol{\beta}^T(u_i, v_i) \mathbf{x}_i \right) \mathbf{x}_i}{\left(1 + \theta_i \exp \left(\boldsymbol{\beta}^T(u_i, v_i) \mathbf{x}_i \right) \right)^2} \right]$$

Turunan-turunan parsial dari persamaan log-likelihood di atas, jika dinyatakan dalam bentuk matriks menjadi:

$$\mathbf{g}^T(\boldsymbol{\beta}_{(m)}) = \left[\frac{\partial \ln L(.)}{\partial \theta}, \frac{\partial \ln L(.)}{\partial \beta_0}, \frac{\partial \ln L(.)}{\partial \beta_1}, \dots, \frac{\partial \ln L(.)}{\partial \beta_p} \right]$$

Proses estimasi parameter koefisien regresi diperoleh melalui metode iterasi numerik Newton Raphson. Metode Newton Raphson digunakan untuk menemukan solusi dari fungsi *log-likelihood* sehingga diperoleh nilai yang cukup konvergen untuk dijadikan sebagai estimasi bagi masing-masing parameter. Algoritma metode Newton Raphson sebagai berikut:

1. Menentukan nilai estimasi awal parameter

$$\hat{\beta}_{(0)} = [\theta_0 \quad \beta_{00} \quad \cdots \quad \beta_{p0}], \text{ iterasi pada saat } m=0$$

2. Membentuk vektor g

$$\mathbf{g}^T(\hat{\beta}_{(m)})_{(p+1)} = \left[\frac{\partial \ln L(.)}{\partial \theta}, \frac{\partial \ln L(.)}{\partial \beta_0}, \frac{\partial \ln L(.)}{\partial \beta_1}, \dots, \frac{\partial \ln L(.)}{\partial \beta_p} \right]_{\beta = \hat{\beta}_{(m)}}$$

dengan p adalah banyaknya parameter yang diestimasi.

3. Membentuk matriks Hessian \mathbf{H} yang elemennya adalah:

$$\mathbf{H}(\hat{\beta}_{(m)})_{(p+1) \times (p+1)} = \begin{bmatrix} \frac{\partial^2 \ln L(.)}{\partial \theta_i^2} & \frac{\partial^2 \ln L(.)}{\partial \theta_1 \partial \beta_0} & \cdots & \frac{\partial^2 \ln L(.)}{\partial \theta_1 \partial \beta_p} \\ \frac{\partial^2 \ln L(.)}{\partial \theta_1 \partial \beta_0} & \frac{\partial^2 \ln L(.)}{\partial \beta_0^2} & \cdots & \frac{\partial^2 \ln L(.)}{\partial \beta_0 \partial \beta_p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \ln L(.)}{\partial \theta_1 \partial \beta_p} & \frac{\partial^2 \ln L(.)}{\partial \beta_0 \partial \beta_p} & \cdots & \frac{\partial^2 \ln L(.)}{\partial \beta_p^2} \end{bmatrix}_{\beta = \hat{\beta}_{(m)}}$$

4. Melakukan iterasi mulai dari $m=0$ pada persamaan:

$$\hat{\beta}_{(m+1)} = \hat{\beta}_{(m)} - \mathbf{H}^{-1}(\hat{\beta}_{(m)}) \mathbf{g}(\hat{\beta}_{(m)})$$

5. Proses iterasi dapat dihentikan ketika nilai estimasi yang diperoleh sudah konvergen ke suatu nilai, atau

$$\hat{\beta}_{(m+1)} \approx \hat{\beta}_{(m)}$$

6. Jika belum mencapai penaksir parameter yang konvergen, maka pada langkah ke-2 hingga konvergen. Penaksir parameter yang konvergen diperoleh jika $\|\hat{\beta}_{(m+1)} - \hat{\beta}_{(m)}\| < \varepsilon$, ε merupakan bilangan yang sangat kecil.

2.7.2 Pengujian Parameter Model GWNBR

Pengujian signifikan parameter model GWNBR terdiri dari uji serentak dan parsial. Uji signifikansi secara serentak dengan menggunakan *Maximum Likelihood Ratio Test* (MLRT) dengan hipotesis sebagai berikut.

$$H_0 : \beta_1(u_i, v_i) = \beta_2(u_i, v_i) = \dots = \beta_p(u_i, v_i) = 0 \quad ; i = 1, 2, \dots, n$$

$$H_1 : \text{minimal ada satu } \beta_k(u_i, v_i) \neq 0; k = 1, 2, \dots, p \quad ; i = 1, 2, \dots, n$$

Statistik Uji :

$$D(\hat{\beta}) = 2(\ln L(\hat{\Omega}) - \ln L(\hat{\omega})) \quad (2.25)$$

Dengan $L(\hat{\omega})$ dan $L(\hat{\Omega})$ sebagai berikut.

$$L(\hat{\omega}) = \prod_{i=1}^n \left(\prod_{r=0}^{y_i-1} (r + \hat{\theta}_i^{-1}) \right) \frac{1}{(y_i!)} \left(\frac{1}{1 + \hat{\theta}_i \hat{\mu}_i} \right)^{1/\hat{\theta}_i} \left(\frac{\hat{\theta}_i \hat{\mu}_i}{1 + \hat{\theta}_i \hat{\mu}_i} \right)$$

$$\ln L(\hat{\omega}) = \sum_{i=1}^n \left[\left(\sum_{r=0}^{y_i-1} \ln(r + \hat{\theta}_i^{-1}) \right) - \ln(y_i!) + y_i \ln \hat{\theta}_i \hat{\mu}_i - (y_i + \hat{\theta}_i^{-1}) \ln(1 + \hat{\theta}_i \hat{\mu}_i) \right]$$

dengan $\hat{\mu}_i = \exp(\hat{\beta}_{0(u_i, v_i)})$

$$L(\hat{\Omega}) = \prod_{i=1}^n \left(\prod_{r=0}^{y_i-1} (r + \hat{\theta}_i^{-1}) \right) \frac{1}{(y_i!)} \left(\frac{1}{1 + \hat{\theta}_i \hat{\mu}_i} \right)^{1/\hat{\theta}_i} \left(\frac{\hat{\theta}_i \hat{\mu}_i}{1 + \hat{\theta}_i \hat{\mu}_i} \right)$$

$$\ln L(\hat{\Omega}) = \sum_{i=1}^n \left[\left(\sum_{r=0}^{y_i-1} \ln(r + \hat{\theta}_i^{-1}) \right) - \ln(y_i!) + y_i \ln \hat{\theta}_i \hat{\mu}_i - (y_i + \hat{\theta}_i^{-1}) \ln(1 + \hat{\theta}_i \hat{\mu}_i) \right]$$

dengan $\hat{\mu}_i = \exp\left(\sum_{l=0}^P \hat{\beta}_{l(u_i, v_i)} x_{il}\right)$

Tolak H_0 jika statistik uji $D(\hat{\beta}) > \chi^2(\alpha; p)$

Pengujian signifikansi secara parsial dilakukan untuk mengetahui parameter mana saja yang signifikan terhadap variabel respon pada tiap-tiap lokasi dengan hipotesis sebagai berikut.

$$H_0 : \beta_k(u_i, v_i) = 0$$

$$H_1 : \beta_k(u_i, v_i) \neq 0; k = 1, 2, \dots, p \quad ; i = 1, 2, \dots, n$$

Statistik Uji:

$$Z_{hit} = \frac{\hat{\beta}_k(u_i, v_i)}{SE(\hat{\beta}_k(u_i, v_i))} \quad (2.26)$$

H_0 ditolak jika $|Z_{hit}| > Z_{(u/2)}$. Tolak H_0 artinya bahwa parameter tersebut berpengaruh signifikan terhadap variabel respon di tiap lokasi (Ricardo dan Carvalho, 2013).

2.7.3 Penentuan *Bandwidth* dan Pembobot Optimum

Secara teoritis *bandwidth* merupakan luasan dengan radius b dari titik pusat lokasi yang digunakan sebagai dasar menentukan bobot setiap pengamatan terhadap model regresi pada lokasi tersebut. Pengamatan-pengamatan yang terletak di dalam radius b masih dianggap berpengaruh terhadap model pada lokasi tersebut sehingga akan diberi bobot tergantung pada fungsi yang digunakan. Nilai *bandwidth* yang sangat kecil menyebabkan varians semakin besar. Hal ini dikarenakan jika nilai *bandwidth* sangat kecil maka akan semakin sedikit pengamatan yang berada dalam radius b , sehingga model yang diperoleh akan sangat kasar (*under smoothing*). Sebaliknya nilai *bandwidth* yang besar akan menimbulkan bias yang semakin besar karena semakin banyak pengamatan yang berada dalam radius b , sehingga model yang diperoleh akan terlampaui halus (*over smoothing*).

Pemilihan *bandwidth* optimum menjadi sangat penting karena akan mempengaruhi ketepatan model terhadap data, yaitu mengatur varians dan bias dari model. Penentuan *bandwidth* optimum dilakukan menggunakan metode *Cross Validation* (CV) sebagai berikut.

$$CV(b) = \sum_{i=1}^n (y_i - \hat{y}_{\neq i}(b))^2 \quad (2.27)$$

$\hat{y}_{\neq i}(b)$ merupakan nilai penaksiran untuk pengamatan ke- i yang dimana saat pemodelannya pengamatan ke- i dihilangkan dari proses penaksiran.

Dalam proses penaksiran parameter model GWNBR di suatu titik (u_i, v_i) dibutuhkan adanya pembobot spasial dimana pembobot yang digunakan adalah sebagai berikut.

Fungsi Kernel *Adaptive Bisquare*:

$$w_{j(u_i, v_i)} = \begin{cases} \left(1 - \left(\frac{d_{ij}}{b_i}\right)^2\right)^2 & \text{untuk } d_{ij} \leq b_i \\ 0 & \text{untuk } d_{ij} > b_i \end{cases} \quad (2.28)$$

dengan $d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2}$ adalah jarak *Euclidean* antara lokasi (u_i, v_i) dan lokasi (u_j, v_j) dan b_i adalah nilai *bandwidth* optimum pada tiap lokasi.

2.8 Penyakit Diare

Menurut Laporan Profil Kesehatan Kota Surabaya (2013) secara klinis penyakit diare adalah penyakit yang ditandai dengan perubahan bentuk dan konsisten tinja melembek sampai mencair serta bertambahnya frekuensi berak lebih dari biasanya (umumnya tiga atau lebih dalam sehari). Penyakit ini dapat digolongkan penyakit ringan, tetapi jika terjadi secara mendadak dan tidak mendapatkan perawatan yang tepat maka diare dapat berakibat fatal misalnya dehidrasi berat yang bisa menyebabkan kematian. Diare dapat menular melalui kuman, bakteri dan virus. Selain itu kondisi iklim yang tidak menentu, kebersihan lingkungan seperti terjadi banjir, sanitasi yang buruk dan juga kurangnya kesadaran masyarakat akan kesehatan turut berperan dalam peningkatan kasus diare. Diare bukanlah penyakit yang datang dengan sendirinya, biasanya ada yang menjadi pemicu terjadinya diare. Secara umum diare kebanyakan disebabkan oleh beberapa infeksi virus tetapi juga seringkali akibat dari racun bakteri. Faktor-faktor yang mempengaruhi terjadinya diare adalah sarana sanitasi, air bersih, lingkungan dan fasilitas kesehatan.

2.9 Faktor-Faktor Mempengaruhi Penyakit Diare

Banyak kajian atau penelitian untuk mengetahui faktor-faktor yang berpengaruh terhadap kejadian diare dan permasalahannya. Silva (2008) melakukan penelitian untuk

mengetahui gambaran pengaruh penyediaan air, jamban keluarga, pengelolaan sampah, sanitasi makanan, fasilitas sanitasi dan pelayanan kesehatan terhadap kejadian diare di Desa Klop Sepuluh, Kecamatan Sukodono, Kabupaten Sidoarjo. Hasil penelitian menunjukkan sebagai berikut.

- a. Sumber air : Kelompok pemakai air dari sumur gali, tidak memasak air sebelum diminum dan memiliki jarak sumur dengan rumah kurang dari 10 m eter memiliki proporsi kejadian diare lebih besar.
- b. Jamban keluarga : Kelompok yang tidak memiliki jamban, tidak memanfaatkan jamban, memiliki kebiasaan untuk buang air besar di sungai dan memiliki jamban yang tidak bersih memiliki proporsi kejadian diare lebih besar.
- c. Pengelolaan sampah : Kelompok yang tidak memiliki tempat sampah, memiliki kebiasaan membuang sampah di belakang rumah dan depan rumah, tempat sampah dengan keadaan tidak tertutup akan menimbulkan lalat yang banyak sehingga memiliki proporsi kejadian diare lebih besar.
- d. Sanitasi makanan : Kelompok yang memiliki kebiasaan menyajikan makanan tanpa tudung saji memiliki proporsi kejadian diare lebih besar.
- e. Fasilitas kesehatan : Kelompok tidak mencuci tangan dengan sabun sebelum makan memiliki proporsi kejadian diare lebih besar.
- f. Pelayanan kesehatan : Kelompok yang tidak paham tentang penyebab diare, tidak pernah mendapat penyuluhan mengenai penyebab, tidak tahu tentang faktor sanitasi, tidak paham tentang hubungan faktor sanitasi dengan penyakit diare memiliki proporsi kejadian diare lebih besar.

Pemerintah telah memberikan upaya dan perhatian dalam bidang sanitasi dengan menetapkan *open defecation free* (ODF) dan peningkatan perilaku hidup bersih dan sehat (PHBS). ODF merupakan sebuah kondisi dimana seluruh individu di suatu daerah tidak lagi melakukan buang air bersih sembarangan dan menggunakan tempat sampah sehat. Melalui program tersebut

diharapkan mampu meningkatkan kualitas dan perilaku masyarakat dalam bidang sanitasi. Selain itu penyakit ditularkan oleh bakteri atau virus melalui kondisi lingkungan. Oleh karena itu jika kondisi lingkungan semakin padat dan termasuk penduduk miskin maka lingkungan sekitar menjadi kumuh dan bakteri maupun virus dari penyakit menular lebih mudah tumbuh.

Pemodelan kejadian diare juga terus berkembang, dimana efek ruang atau lokasi dalam spasial juga berpengaruh. Seperti yang dilakukan oleh Kazembe (2009) melakukan pemodelan dengan regresi logistic dan model spasial pada mordibitas demam dan diare anak-anak di Negara Malawi. Hasil penelitian menunjukkan bahwa daerah bagian tertentu memiliki spesifikasi pengaruh yang tinggi dan prevalensi demam dan diare sangat dipengaruhi oleh individu, keluarga dan faktor resiko masyarakat.

Arumsari (2011) melakukan pemodelan *Ordinary Least Square* (OLS), regresi Poisson dan *Geographically Weighted Poisson Regression* (GWPR) untuk mengetahui faktor-faktor yang berpengaruh pada kejadian diare di Kabupaten Tuban Jawa Timur. GWPR merupakan model spasial dengan pendekatan titik. Variabel yang signifikan berpengaruh dari ketiga model tersebut adalah kesediaan fasilitas air minum sendiri dan jarak rumah dengan tempat pembuangan akhir tinja (kurang dari 10 meter). Dengan membandingkan nilai AIC pada ketiga model didapatkan model GWPR adalah yang terbaik.

Beberapa penelitian tersebut menjadi referensi pendukung untuk pengembangan pemodelan spasial pada penelitian tentang kejadian diare. Sehingga diketahui bahwa pemodelan untuk mengetahui faktor-faktor yang berpengaruh terhadap kejadian diare dengan mengikutsertakan pengaruh spasial yang juga dianggap sangat penting.

“Halaman ini sengaja dikosongkan”

BAB III METODOLOGI PENELITIAN

3.1 Sumber Data

Sumber data pada penelitian ini menggunakan data sekunder yaitu mengenai faktor-faktor yang mempengaruhi terjadinya penyakit diare di Kota Surabaya yang diperoleh dari Profil Kesehatan Kota Surabaya 2013. Unit analisis yang digunakan adalah 31 kecamatan di Kota Surabaya.

3.2 Variabel Penelitian

Variabel penelitian dalam penelitian ini adalah sebagai berikut.

Tabel 3.1 Variabel Penelitian

Variabel	Keterangan	Tipe Variabel
Y	Jumlah kejadian penyakit diare (Y)	Diskrit
X ₁	Persentase Rumah Sehat	Kontinyu
X ₂	Persentase Rumah Tangga (RT) yang memiliki jamban sehat	Kontinyu
X ₃	Persentase RT yang memiliki tempat sampah sehat	Kontinyu
X ₄	Persentase rumah tangga miskin	Kontinyu
X ₅	Persentase rumah tangga ber Perilaku Hidup Bersih dan Sehat (PHBS)	Kontinyu
X ₆	Kepadatan penduduk	Kontinyu
X ₇	Rasio sarana kesehatan puskesmas	Kontinyu
X ₈	Rasio jdokter umum	Kontinyu
u_i	Lintang (longitude) kecamatan ke- <i>i</i>	
v_i	Bujur (latitude) kecamatan ke- <i>i</i>	

Definisi operasional variabel penelitian dapat dijelaskan sebagai berikut.

1. Persentase Rumah Sehat (X₁) merupakan hasil bagi rumah sehat terhadap seluruh rumah tangga yang diperiksa rumahnya

di setiap kecamatan Kota Surabaya dikalikan dengan 100%. Rumah sehat adalah bangunan rumah tinggal yang memenuhi syarat kesehatan, yaitu rumah memiliki jamban sehat, sarana air bersih, tempat pembuangan sampah, sarana pembuangan air limbah, ventilasi rumah yang baik, kepadatan hunian rumah yang sesuai dan lantai rumah yang tidak terbuat dari tanah.

2. Persentase Rumah Tangga (RT) yang memiliki jamban sehat (X_2) merupakan hasil bagi dari jumlah RT yang memiliki jamban sehat terhadap jumlah RT yang ada di setiap kecamatan Kota Surabaya dikalikan 100%. Dalam hal ini jamban sehat adalah jamban yang bersih dan tidak digunakan untuk umum.
3. Persentase RT yang memiliki tempat sampah sehat (X_3) merupakan hasil bagi RT yang memiliki tempat sampah sehat terhadap seluruh rumah tangga yang diperiksa tempat sampahnya di setiap kecamatan Kota Surabaya dikalikan dengan 100%. Tempat sampah sehat merupakan tempat pembuangan sampah yang konstruksinya memenuhi syarat-syarat kesehatan (ketentuan program).
4. Persentase RT Miskin (X_4) merupakan hasil bagi antara jumlah penduduk miskin terhadap jumlah penduduk seluruhnya di setiap kecamatan Kota Surabaya dikalikan dengan 100%. Pengukuran kemiskinan yang dihitung oleh Badan Pusat Statistik (BPS) dengan cara menetapkan standar kebutuhan minimum baik untuk makanan maupun minuman yang harus dipenuhi seseorang untuk hidup layak.
5. Persentase rumah tangga ber Perilaku Hidup Bersih dan Sehat (PHBS) (X_5) merupakan hasil bagi dari jumlah rumah tangga yang ber PHBS terhadap jumlah rumah tangga yang dipantau di setiap kecamatan Kota Surabaya dikalikan dengan 100%. PHBS adalah upaya untuk memberdayakan anggota rumah tangga agar tahu, mau serta berperan aktif dalam gerakan kesehatan di masyarakat.
6. Kepadatan penduduk (X_6) merupakan hasil bagi dari jumlah penduduk terhadap luas wilayah di setiap kecamatan Kota

Surabaya dalam satuan (orang/km²). Kepadatan penduduk yang tidak seimbang dengan luas wilayah akan memunculkan area yang sifatnya kumuh tidak beraturan sehingga dapat menyebabkan masalah kesehatan masyarakat.

7. Rasio sarana kesehatan puskesmas (X_7) merupakan hasil bagi jumlah puskesmas terhadap jumlah penduduk di setiap kecamatan Kota Surabaya.
8. Rasio dokter umum (X_8) merupakan hasil bagi jumlah dokter umum di Puskesmas setiap kecamatan Kota Surabaya terhadap jumlah penduduk.

Tabel 3.2 Struktur Data Penelitian

Kecamatan	Y	X_1	X_2	...	X_8	u_i	v_i
1	y_1	$x_{1;1}$	$x_{2;1}$...	$X_{8;1}$	u_1	v_1
2	y_2	$x_{1;2}$	$x_{2;2}$...	$X_{8;2}$	u_2	v_2
3	y_3	$x_{1;3}$	$x_{2;3}$...	$X_{8;3}$	u_3	v_3
\vdots	\vdots	\vdots	\vdots	\ddots	\vdots	\vdots	\vdots
31	y_{31}	$x_{1;31}$	$x_{2;31}$...	$X_{8;31}$	u_{31}	v_{31}

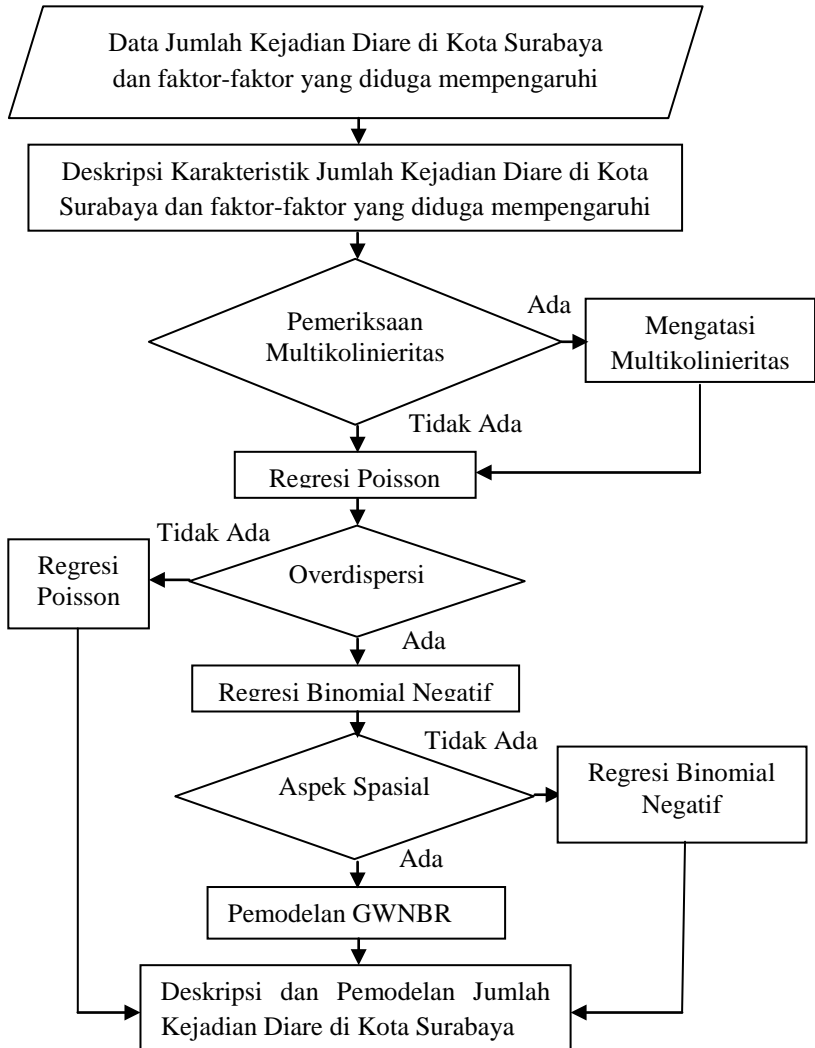
3.3 Langkah Analisis

Langkah-Langkah dalam menganalisis penelitian ini adalah dengan tahapan sebagai berikut.

1. Melakukan analisis statistika deskriptif untuk mengetahui deskripsi karakteristik kejadian diare beserta faktor-faktor yang diduga mempengaruhi di setiap kecamatan Kota Surabaya.
2. Memodelkan faktor-faktor yang berpengaruh terhadap kejadian diare di Kota Surabaya menggunakan GWNBR.
 - a. Pemeriksaan adanya multikolinieritas antar variabel prediktor menggunakan nilai VIF (*Variance Inflation Factor*), mengatasi adanya multikolinieritas dengan mengeluarkan variabel prediktor satu per satu mulai dari yang memiliki nilai VIF paling besar.
 - b. Menganalisis model regresi Poisson dengan langkah sebagai berikut.

- i. Menaksir parameter model regresi Poisson.
- ii. Menguji signifikansi parameter regresi Poisson secara serentak dan parsial.
- iii. Melakukan pemeriksaan dispersi model regresi Poisson yaitu dengan cara membagi nilai devians dengan derajat bebas.
- c. Menganalisis model regresi Binomial Negatif dengan langkah sebagai berikut.
 - i. Menaksir parameter model regresi Binomial Negatif.
 - ii. Menguji signifikansi parameter regresi Binomial Negatif secara serentak dan parsial.
- d. Menguji aspek data spasial, yang meliputi uji *BP* untuk melihat heterogenitas spasial data dan uji Moran's I untuk menguji dependensi spasial data.
- e. Menganalisis model GWNBR dengan langkah sebagai berikut.
 - i. Menghitung jarak *euclidean* antar lokasi pengamatan berdasarkan posisi geografis. Jarak euclidean antara lokasi i yang terletak pada koordinat (u_i, v_i) terhadap lokasi j yang terletak pada koordinat (u_j, v_j) .
 - ii. Menentukan *bandwidth* optimum untuk setiap lokasi pengamatan dengan menggunakan Cross Validation (CV).
 - iii. Menghitung matriks pembobot dengan fungsi kernel *Adaptive Bisquare*.
 - iv. Menaksir parameter model GWNBR.
 - v. Menguji signifikansi parameter model GWNBR secara serentak dan parsial.
 - vi. Menginterpretasikan hasil model GWNBR.

Tahapan langkah analisis dapat dilihat pada diagram alir Gambar 3.1 sebagai berikut.



Gambar 3.1 Diagram Alir Langkah Analisis

“Halaman ini sengaja dikosongkan”

BAB IV ANALISIS DAN PEMBAHASAN

Pada bab IV ini membahas deskripsi jumlah kejadian diare beserta faktor yang diduga mempengaruhinya di Kota Surabaya. Selanjutnya dilakukan pemetaan kecamatan dengan menggunakan peta tematik dan pemodelan menggunakan metode analisis *Geographically Weighted Negative Binomial Regression* untuk mengetahui faktor-faktor yang mempengaruhi kejadian diare di masing-masing kecamatan Kota Surabaya.

4.1 Deskripsi Jumlah Kejadian Diare Beserta Faktor-Faktor yang diduga Mempengaruhinya di Kota Surabaya

Deskripsi karakteristik mengenai jumlah kejadian diare beserta faktor-faktor yang diduga mempengaruhinya akan disajikan pada Tabel 4.1 sebagai berikut.

Tabel 4.1 Karakteristik Jumlah Kejadian Diare Beserta Faktor-Faktor yang diduga Mempengaruhinya

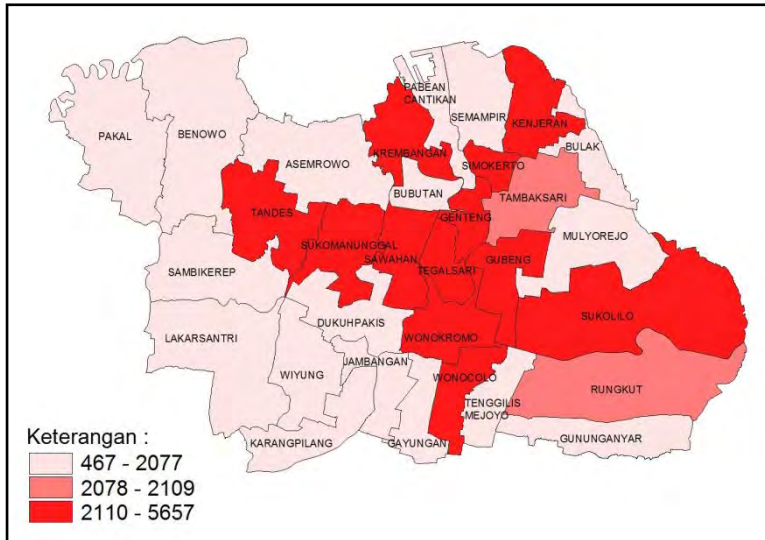
Variabel	Rata-Rata	Standar Deviasi	Minimum	Maksimum
Jumlah kejadian diare (Y)	2093	1233	467	5657
Persentase rumah sehat (X_1)	80,93	9,53	55,43	100
Persentase RT memiliki yang jamban sehat (X_2)	96,80	7,11	59,89	100
Persentase RT yang memiliki tempat sampah sehat (X_3)	85,67	16,17	35,35	100
Persentase rumah tangga miskin (X_4)	10,64	5,480	4,870	28
Persentase rumah tangga ber-PHBS (X_5)	69,81	14,84	34,26	91,61
Kepadatan penduduk (X_6)	11344	7168	2190	31194
Rasio sarana kesehatan puskesmas (X_7)	0,668	0,273	0,082	1,446
Rasio dokter umum (X_8)	0,238	0,088	0,027	0,527

Berdasarkan Tabel 4.1 diketahui bahwa rata-rata kejadian diare di Kota Surabaya pada tahun 2013 adalah sebesar 2093 kejadian, dengan kejadian diare minimum 467 kejadian yaitu di Kecamatan Pakal sedangkan kejadian maksimum sebanyak 5657 kejadian untuk Kecamatan Sawahan. Surabaya merupakan kota metropolitan dengan segala aktivitas pembangunan infrastruktur yang semakin meningkat, sehingga untuk persentase rumah sehat, RT yang memiliki jamban sehat dan tempat sampah sehat sudah ada beberapa kecamatan yang menunjukkan 100 persen atau mencapai batas maksimum. Sedangkan untuk persentase RT miskin di Kota Surabaya minimum sebesar 4,87 persen yaitu Kecamatan Tenggilis Mejoyo dan maksimum di Kecamatan Semampir yaitu sebesar 28 persen dengan keragaman data di Kota Surabaya sebesar 5,48. Jadi walaupun Kota Surabaya semakin meningkat dalam hal upaya pengembangannya masih saja ditemui adanya rumah tangga miskin.

Persentase RT yang berperilaku hidup bersih dan sehat memiliki rata-rata sebesar 69,81 persen dengan nilai maksimum 91,61 persen dan minimum 34,26 persen. Selain itu Kota Surabaya juga terkenal dengan kepadatan penduduknya, kecamatan yang paling tinggi kepadatan penduduk adalah Kecamatan Simokerto yaitu sebanyak 3131193,82 jiwa per km². Dalam upaya pembangunan kesehatan Kota Surabaya memiliki rata-rata rasio dokter umum sebesar 0,668 dan sarana kesehatan puskesmas sebesar 0,238. Kota Surabaya terus melakukan upaya pengembangan terutama di bidang kesehatan selain upaya pembangunan infrastrukturnya. Hal tersebut diupayakan untuk mengurangi kejadian penyakit khususnya penyakit yang dapat menular melalui kondisi sanitasi lingkungan yang buruk.

4.1.1 Jumlah Kejadian Diare Tahun 2013 di Kota Surabaya

Berikut akan dibahas mengenai kejadian diare di Kota Surabaya pada tahun 2013 dengan menggunakan peta tematik.



Gambar 4.1 Persebaran Jumlah Kejadian Diare Kota Surabaya

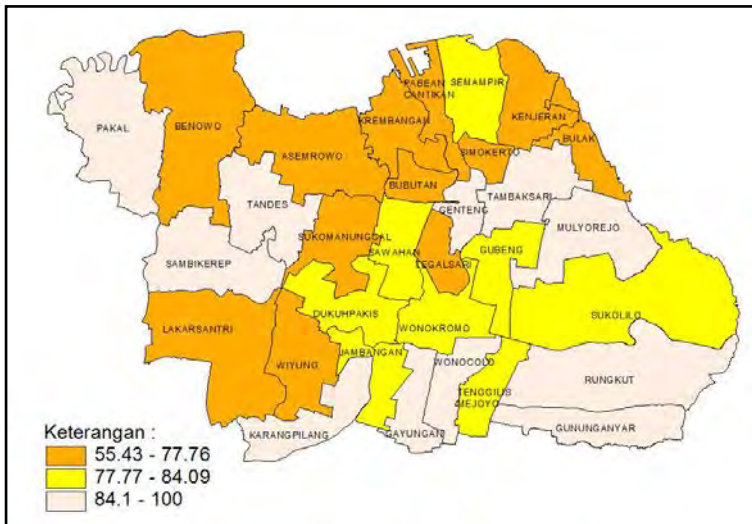
Jumlah kejadian diare di Kota Surabaya diketahui tinggi adalah Kecamatan Kenjeran yaitu 5602 kejadian dan Kecamatan Sawahan sebanyak 5657 kejadian. Hal tersebut dikarenakan letak kondisi Kecamatan Kenjeran merupakan kawasan pantai dan dikenal sebagai kawasan kumuh sehingga dapat menyebarkan virus yang menyebabkan terjangkitnya penyakit diare. Selain itu banyak orang yang tinggal di kawasan tersebut tidak menerapkan perilaku hidup bersih dan sehat sehingga mudah memicu terjadinya lingkungan yang kotor dan tidak sehat bahkan masih ada yang menggunakan jamban secara umum.

Kejadian diare tertinggi ada di Kecamatan Sawahan yang juga merupakan lingkungan padat penduduk sehingga dapat menyebarkan virus penyakit menular seperti halnya penyakit diare. Wilayah yang termasuk kejadian diare antara 2078 hingga 2109 kejadian adalah Kecamatan Rungkut dan Tambaksari. Wilayah yang termasuk kejadian diare antara 467 hingga 2077 kejadian antara lain Kecamatan Pakal, Asemrowo, Gunung

Anyar. Hal tersebut dikarenakan untuk wilayah tersebut merupakan wilayah yang tidak terlalu padat penduduk sehingga dapat meminimalisir penyebaran kuman atau virus. Selain itu juga wilayah dengan kawasan yang merupakan pembangunan Kota Surabaya, jadi untuk kesadaran perilaku hidup bersih dan sehat juga tinggi.

4.1.2 Persentase Rumah Sehat

Persentase rumah sehat merupakan rumah yang diperiksa untuk kebersihan dan dapat dikatakan sehat jika telah memenuhi kriteria yang telah ditentukan. Untuk mengetahui persebaran persentase rumah sehat di Kota Surabaya dapat dilihat pada Gambar 4.2 berikut.



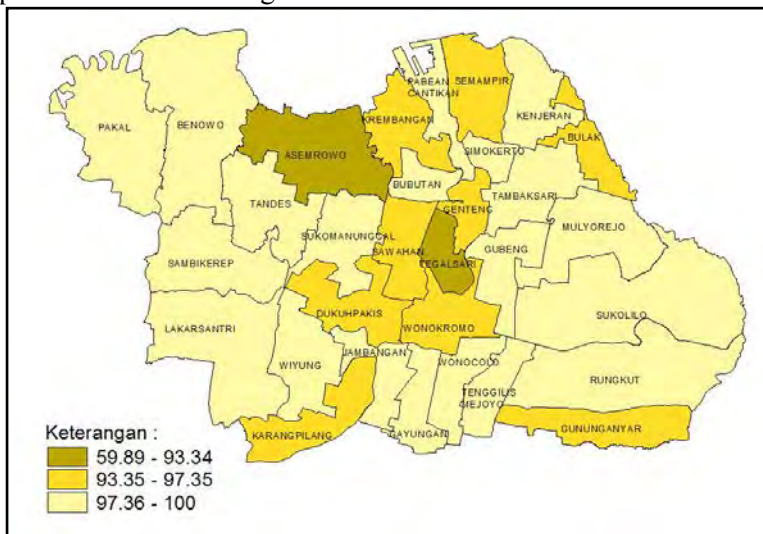
Gambar 4.2 Persebaran Persentase Rumah Sehat

Persentase rumah tangga yang memiliki kriteria rumah sehat dengan nilai maksimum 100 persen adalah Kecamatan Gunung Anyar. Hal tersebut dikarenakan kondisi yang tidak terlalu padat dengan penduduk. Salah satu kriteria rumah sehat adalah rumah yang memiliki ventilasi udara. Jadi menurut Dinas

kesehatan kota Surabaya seluruh rumah tangga Kecamatan Gunung Anyar telah memenuhi kriteria rumah sehat. Sebaliknya Kecamatan Lakarsantri merupakan kecamatan terendah yang memiliki persentase rumah sehat yaitu sebesar 55,43 persen. Hal tersebut terlihat bahwa Kecamatan Benowo merupakan kawasan barat Kota Surabaya yang masih tergolong kurang perkembangannya terutama dalam bidang kesehatan.

4.1.3 Persentase Rumah Tangga Memiliki Jamban Sehat

Dalam upaya pembangunan kesehatan di Kota Surabaya akan lebih diinginkan kawasan yang sehat misalnya keadaan pada saat MCK (Mandi Cuci Kakus). Oleh karena itu akan diketahui persebaran persentase rumah tangga yang memiliki jamban sehat pada Gambar 4.3 sebagai berikut.



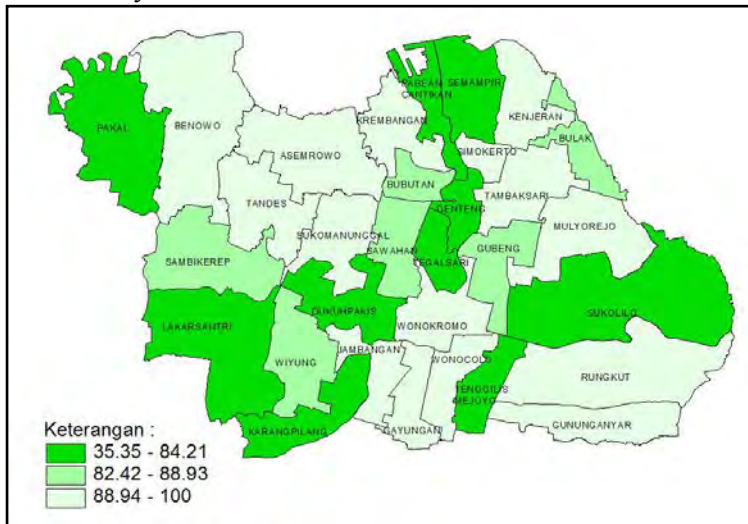
Gambar 4.3 Persebaran Persentase RT Memiliki Jamban Sehat

Kecamatan Asemrowo merupakan kecamatan yang tertinggal dalam hal kepemilikan jamban sehat di Kota Surabaya yaitu hanya sebesar 59,89 persen RT yang memiliki jamban sehat. Jamban sehat adalah salah satu hal penting untuk

menciptakan kesehatan, karena dengan jamban yang tidak sehat atau digunakan secara umum maka akan dapat memicu timbulnya bakteri atau virus yang dapat menyebabkan penyakit. Selebihnya untuk kecamatan yang lain sudah diatas 90 persen rumah tangga memiliki jamban sehat. Hal tersebut dikarenakan upaya pembangunan kesehatan telah digalakkan untuk tercapainya masyarakat yang sehat dan sejahtera tanpa penyakit.

4.1.4 Persentase Rumah Tangga Memiliki Tempat Sampah Sehat

Persebaran persentase rumah tangga yang memiliki tempat sampah sehat dapat dilihat pada Gambar 4.4 tempat sampah sehat merupakan tempat pembuangan sampah yang kontruksinya memenuhi syarat kesehatan.



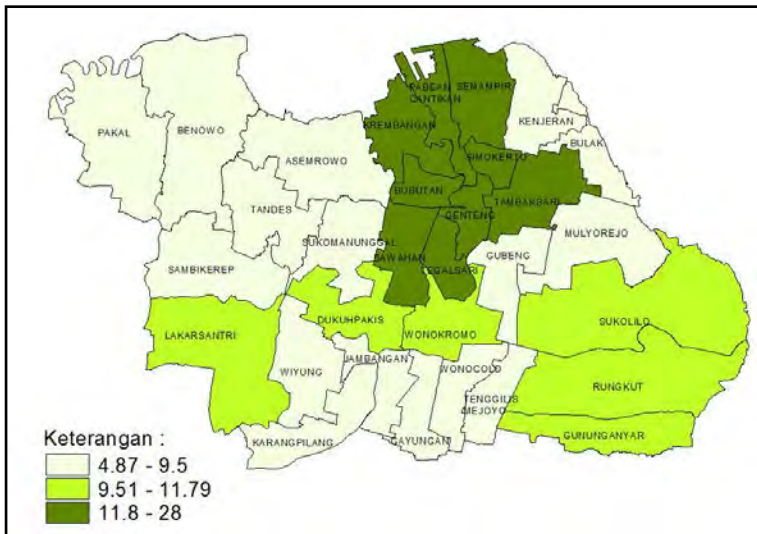
Gambar 4.4 Persebaran Persentase RT Memiliki Tempat Sampah Sehat

Kecamatan Pakal dan Karangpilang merupakan kawasan yang masih rendah memiliki tempat sampah sehat. Namun sebagian besar RT di tiap kecamatan Kota Surabaya telah memiliki tempat sampah sehat yaitu diatas 88,93 persen

diantaranya adalah Kecamatan Tambaksari, Mulyorejo dan Rungkut. Rumah tangga yang memiliki tempat sampah sehat menunjukkan bahwa resiko terjangkit penyakit menular akan berkurang karena memiliki lingkungan yang bersih dan tidak kumuh akibat sampah.

4.1.5 Persentase Rumah Tangga Miskin

Kota Surabaya tergolong kota yang besar dan maju dalam upaya pembangunannya, namun tak terlepas dari adanya kondisi rumah tangga miskin. Berikut akan dibahas mengenai persebaran persentase rumah tangga miskin di Kota Surabaya pada Gambar 4.5 berikut ini.



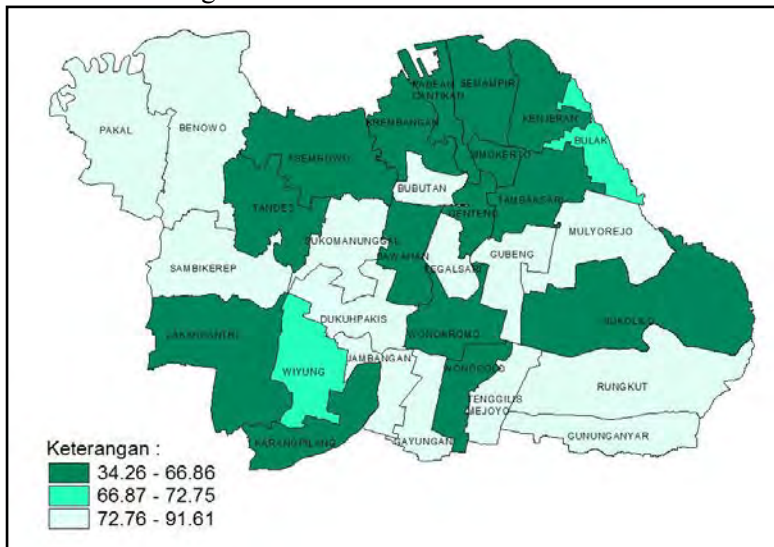
Gambar 4.5 Persebaran Persentase RT Miskin

Persentase rumah tangga miskin kategori tinggi adalah Kecamatan Pabean Cantikan, Semampir dan Simokerto. Wilayah tersebut memiliki persentase rumah tangga miskin diatas 11,8 hingga 28 persen sehingga memungkinkan keadaan yang memicu adanya bakteri atau virus serta terjangkitnya suatu penyakit. Namun sebagian besar persentase rumah tangga miskin di Kota

Surabaya antara 9,51 hingga 11,79 persen diantaranya adalah Kecamatan Sukolilo, Rungkut dan Gunung Anyar. Sedangkan persentase rumah tangga miskin yang rendah yaitu antara 4,87 hingga 9,5 persen adalah diantaranya Kecamatan Gubeng, Mulyorejo dan Tenggilis Mejoyo. Persentase rumah tangga miskin di Kota Surabaya mencapai angka terbesar adalah 28 persen sehingga pemberantasan rumah tangga miskin cukup diupayakan pemerintah agar terwujud masyarakat yang sejahtera.

4.1.6 Persentase Rumah Tangga Perilaku Hidup Bersih dan Sehat

Perilaku hidup bersih dan sehat sangat diperlukan untuk menjaga kesehatan menjadi yang lebih baik. persebaran persentase rumah tangga yang ber-PHBS dapat dilihat pada Gambar 4.6 sebagai berikut.



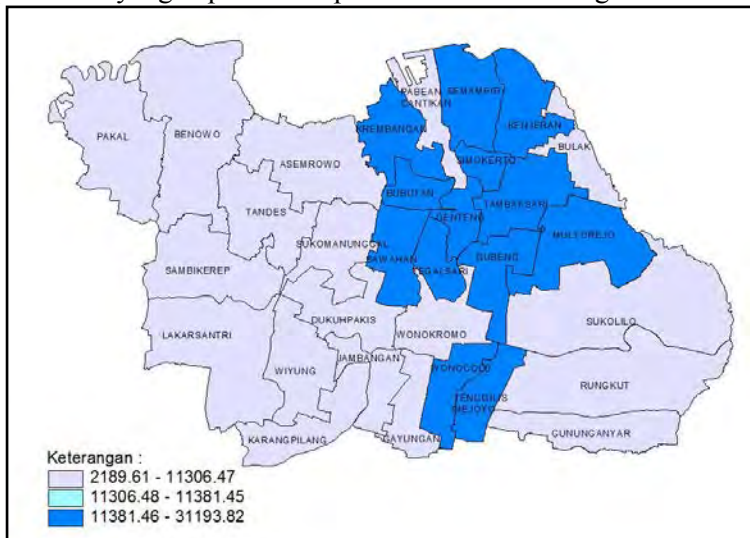
Gambar 4.6 Persebaran RT ber-PHBS

Persentase rumah tangga yang ber-PHBS paling rendah adalah Kecamatan Wonokromo yaitu sebesar 34,26 persen. Untuk

persentase antara 66,87 hingga 72,75 persen adalah Kecamatan Wiyung dan Bulak. Sebagian besar di Kota Surabaya wilayah dengan rumah tangga ber-PHBS memiliki persentase antara 72,76 hingga 91,61 antara lain adalah Kecamatan Rungkut, Gubeng dan Mulyorejo. Rumah tangga ber-PHBS dapat mengurangi resiko terkena penyakit yang dapat menular melalui kondisi lingkungan sekitar misalnya penyakit diare. Hal tersebut dikarenakan perilaku hidup bersih dan sehat memiliki sanitasi yang baik sehingga dapat meminimalisir berkembangnya virus maupun bakteri dari penyakit menular.

4.1.7 Kepadatan Penduduk

Kota Surabaya terkenal dengan kepadatan penduduknya. Oleh karena itu akan dibahas mengenai persebaran kepadatan penduduk yang dapat dilihat pada Gambar 4.7 sebagai berikut.



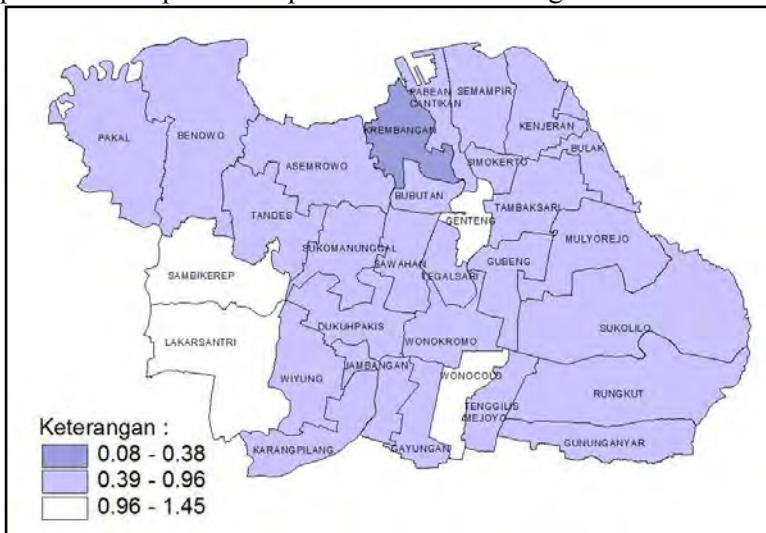
Gambar 4.7 Persebaran Kepadatan Penduduk

Berdasarkan Gambar 4.7 di ketahui bahwa Kecamatan Sawahan, Tegalsari, Bubutan, Simokerto, Semampir, Kenjeran dan Tambaksari merupakan wilayah yang memiliki kepadatan

penduduk diatas 16334,42 jiwa per km². Lingkungan yang padat akan penduduknya dapat menyebabkan lingkungan menjadi kumuh dan tidak sehat. Jadi semakin padat penduduk suatu wilayah maka diduga jumlah kejadian diare juga akan semakin tinggi. Kepadatan penduduk menyebabkan keadaan lingkungan menjadi kumuh dan bisa dijadikan sarang binatang seperti halnya lalat yang juga dapat berkembang menjadi virus atau bakteri penyakit menular.

4.1.8 Rasio Sarana Kesehatan Puskesmas

Rasio sarana kesehatan puskesmas dihitung untuk mengetahui suatu wilayah emmbutuhkan puskesmas tambahan atau tidak. Menurut Dinas Kesehatan, satu puskesmas melayani 30.000 penduduk. Untuk mengetahui persentase persebaran rasio puskesmas dapat dilihat pada Gambar 4.8 sebagai berikut.



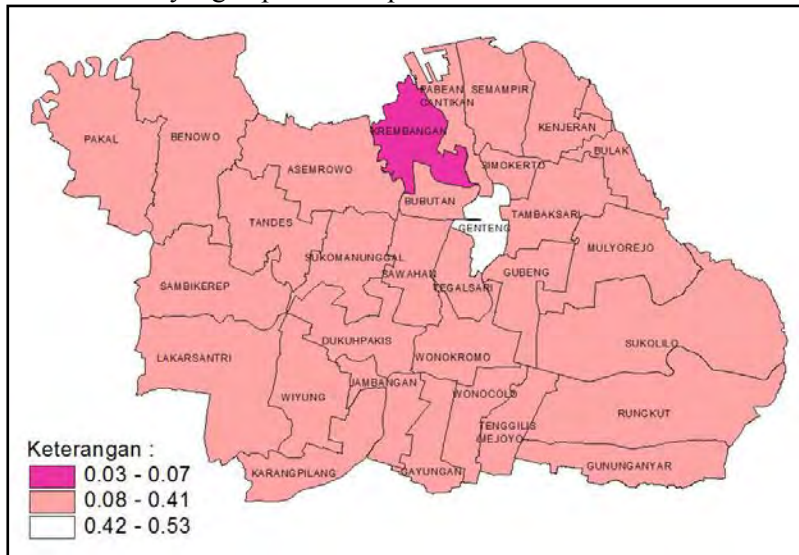
Gambar 4.8 Persebaran Rasio Sarana Kesehatan Puskesmas

Gambar 4.8 menunjukkan bahwa rasio puskesmas terendah yaitu dibawah 0,08 adalah Kecamatan Krembangan. Sedangkan wilayah yang tergolong rasio puskesmas tinggi yaitu diatas 0,97 adalah Kecamatan Sambikerep, Lakarsantri, Wonokromo dan

Genteng. Puskesmas merupakan salah satu sarana kesehatan yang dibangun oleh pemerintah dengan tujuan untuk membantu pelayanan kesehatan masyarakat. Selain itu diharapkan dapat memberikan penyuluhan mengenai pencegahan penyakit.

4.1.9 Rasio Dokter Umum

Rasio sebaran dokter umum terhadap penduduk berdasarkan standar Organisasi Kesehatan Dunia (WHO) pemerintah menargetkan 1 dok ter umum melayani 2.500 penduduk. Berikut akan dibahas mengenai persebaran rasio dokter umum yang dapat dilihat pada Gambar 4.9.



Gambar 4.9 Persebaran Rasio Dokter Umum

Berdasarkan Gambar 4.9 diketahui bahwa rasio dokter umum yang tinggi yaitu 0,53 adalah Kecamatan Genteng. Sedangkan yang paling rendah adalah Kecamatan Krembangan sebesar 0,03 persen. Semakin banyak adanya dokter umum di setiap kecamatan maka dapat menangani penderita diare sehingga dapat mengurangi jumlah kejadian diare di Kota Surabaya. Selain itu juga dokter umum dapat melakukan penyuluhan mengenai

pencegahan terjadinya penyakit diare dengan koordinasi dari pihak terkait misalnya puskesmas. Oleh karena itu rasio dokter umum diharapkan semakin meningkat agar kondisi kesehatan lebih dapat ditingkatkan dan meminimalisir kejadian diare khususnya.

4.2 Pemodelan Jumlah Kejadian Diare di Kota Surabaya

Data yang digunakan dalam bentuk data *count* yaitu jumlah kejadian diare di Kota Surabaya, jadi pembentukkan model menggunakan regresi Poisson. Namun pada penelitian ini diduga adanya kasus overdispersi dan ingin diketahui faktor-faktor yang berpengaruh terhadap jumlah kejadian diare di masing-masing kecamatan Kota Surabaya sehingga metode yang sesuai digunakan adalah *Geographically Weighted negative Binomial Regression* (GWNBR). *Software* yang digunakan dalam analisis pemodelan adalah *software* R.

4.2.1 Pemeriksaan Multikolineritas

Beberapa cara untuk mendeteksi adanya kasus multikolineritas adalah dengan melihat koefisien korelasi antar variabel prediktor dan nilai *Variance Inflation Factor* (VIF).

Tabel 4.2 Koefisien Korelasi Antara Variabel Prediktor

	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇
X ₂	0,185						
X ₃	-0,055	-0,055					
X ₄	-0,217	-0,033	-0,042				
X ₅	0,130	0,301	0,038	-0,298			
X ₆	-0,071	0,181	0,176	0,501	-0,045		
X ₇	0,097	0,009	-0,043	-0,048	-0,117	-0,168	
X ₈	0,156	0,033	-0,009	-0,076	-0,222	-0,155	0,828

Berdasarkan Tabel 4.2 diketahui bahwa nilai koefisien korelasi antar variabel prediktor menunjukkan tidak adanya kasus multikolineritas. Hal tersebut dikarenakan nilai koefisien korelasi tidak mendekati nilai satu. Selanjutnya untuk lebih jelas juga dilakukan dengan melihat nilai VIF yaitu sebagai berikut.

Tabel 4.3 Nilai VIF Masing-Masing Variabel Prediktor

Variabel	VIF
X_1	1,121
X_2	1,219
X_3	1,082
X_4	1,594
X_5	1,357
X_6	1,547
X_7	3,327
X_8	3,572

Jika nilai VIF lebih dari 10 maka dideteksi adanya kasus multikolinieritas. Dapat dilihat pada Tabel 4.3 m enunjukkan bahwa nilai VIF dari masing-masing variabel prediktor memiliki nilai VIF kurang dari 10 sehingga dapat disimpulkan tidak adanya kasus multikolinieritas dalam hal ini.

4.2.2 Pemodelan Regresi Poisson dan Regresi Binomial Negatif

Selanjutnya setelah diketahui tidak adanya multikolinieritas dalam hal ini, maka dapat dilakukan analisis pemodelan regresi Poisson. Berikut ini pada Tabel 4.4 merupakan hasil penaksiran parameter regresi Poisson.

Tabel 4.4 Penaksiran Parameter Model Regresi Poisson

Parameter	Estimasi	Z	$\Pr(Z > Z_{hit})$
$\hat{\beta}_0$	7,04800	78,171	$< 2 \times 10^{-16}*$
$\hat{\beta}_1$	-0,01182	-24,179	$< 2 \times 10^{-16}*$
$\hat{\beta}_2$	0,01920	23,607	$< 2 \times 10^{-16}*$
$\hat{\beta}_3$	0,00136	4,753	$2,01 \times 10^{-16}*$
$\hat{\beta}_4$	-0,03619	-37,728	$< 2 \times 10^{-16}*$
$\hat{\beta}_5$	-0,01443	-46,990	$< 2 \times 10^{-16}*$
$\hat{\beta}_6$	0,00005	68,785	$< 2 \times 10^{-16}*$
$\hat{\beta}_7$	-0,34430	-11,941	$< 2 \times 10^{-16}*$
$\hat{\beta}_8$	2,43400	29,016	$< 2 \times 10^{-16}*$
Devians = 8803,8		Df = 22	

*) signifikan dengan taraf nyata 5%

Pada pemodelan regresi Poisson, pengujian signifikansi parameter secara serentak dilakukan untuk mengetahui ada atau tidaknya pengaruh secara serentak variabel prediktor terhadap variabel respon. Hipotesis yang digunakan dan pembahasannya adalah sebagai berikut.

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = \beta_8 = 0$$

$$H_1 : \text{paling sedikit ada satu } \beta_k \neq 0, k=1,2,\dots,8$$

Berdasarkan hasil analisis pengujian serentak pada regresi Poisson dengan taraf signifikansi 5% didapatkan $\chi^2_{(8;0,05)}$ sebesar 15,507 yang artinya minimal ada satu variabel prediktor yang berpengaruh signifikan terhadap variabel respon. Selanjutnya dilakukan pengujian signifikansi parameter secara parsial untuk mengetahui variabel prediktor mana yang berpengaruh terhadap variabel respon dengan hipotesis sebagai berikut.

$$H_0 : \beta_k = 0 \text{ (variabel ke-} k \text{ tidak signifikan)}$$

$$H_1 : \beta_k \neq 0 \text{ (variabel ke-} k \text{ signifikan)} ; k = 1, 2, \dots, 8$$

Pada hasil analisis regresi Poisson menunjukkan bahwa pengujian parsial dengan taraf signifikansi 5% didapatkan $z_{(0,05/2)}$ sebesar 1,96 yang artinya berdasarkan Tabel 4.4 semua variabel prediktor berpengaruh signifikan terhadap variabel respon. Sehingga didapatkan model regresi Poisson sebagai berikut.

$$\hat{\mu} = \exp \left(\begin{array}{l} 7,048 - 0,0118X_1 + 0,0192X_2 + 0,0013X_3 - 0,0361X_4 - 0,0144X_5 \\ + 0,00004X_6 - 0,3443X_7 + 2,434X_8 \end{array} \right)$$

Selain itu juga dapat diketahui berdasarkan Tabel 4.4 menunjukkan bahwa nilai rasio antara devians dengan derajat bebas dari regresi Poisson adalah lebih besar dari 1 sehingga dapat dideteksi adanya kasus overdispersi. Oleh karena itu untuk mengatasi overdispersi pada kejadian diare di Kota Surabaya maka dapat menggunakan analisis model regresi binomial negatif.

Tabel 4.5 Penaksiran Parameter Model Regresi Binomial Negatif

Parameter	Estimasi	Z	$\Pr(Z > Z_{hit})$
$\hat{\beta}_0$	6,11300	5,257	1,47e-07*
$\hat{\beta}_1$	-0,00723	-0,975	0,32948
$\hat{\beta}_2$	0,02180	2,102	0,03556*
$\hat{\beta}_3$	0,00553	1,287	0,19805
$\hat{\beta}_4$	-0,02391	-1,555	0,11987
$\hat{\beta}_5$	-0,01672	-3,192	0,00141*
$\hat{\beta}_6$	0,00005	3,972	0,00007*
$\hat{\beta}_7$	-0,02544	-0,057	0,95436
$\hat{\beta}_8$	1,53300	1,081	0,27984
Devians = 31,671		Df = 22	

*) signifikan dengan taraf nyata 5%

Pengujian signifikansi parameter secara serentak pada regresi binomial negatif dilakukan untuk mengetahui ada atau tidaknya pengaruh secara serentak variabel prediktor terhadap variabel respon. Hipotesis yang digunakan adalah sebagai berikut.

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = \beta_8 = 0$$

$$H_1 : \text{paling sedikit ada satu } \beta_k \neq 0, k=1,2,\dots,8$$

Hasil analisis pengujian serentak pada regresi binomial negatif dengan taraf signifikansi 5% didapatkan $\chi^2_{(8;0,05)}$ sebesar 15,507 yang artinya minimal ada satu variabel prediktor yang berpengaruh signifikan terhadap variabel respon. Selanjutnya dilakukan pengujian signifikansi parameter secara parsial untuk mengetahui variabel prediktor mana yang berpengaruh terhadap variabel respon dengan hipotesis sebagai berikut.

$$H_0 : \beta_k = 0 \text{ (variabel ke-} k \text{ tidak signifikan)}$$

$$H_1 : \beta_k \neq 0 \text{ (variabel ke-} k \text{ signifikan)} ; k = 1, 2, \dots, 8$$

Pada hasil analisis regresi binomial negatif menunjukkan bahwa pengujian parsial dengan taraf signifikansi 5% didapatkan $z_{(0,05/2)}$ sebesar 1,96 yang artinya berdasarkan Tabel 4.5 variabel yang berpengaruh signifikan terhadap kejadian diare di Kota Surabaya adalah persentase RT memiliki jamban sehat (X_2), persentase RT ber-PHBS (X_5) dan kepadatan penduduk (X_6).

Sehingga didapatkan model regresi binomial negatif adalah sebagai berikut.

$$\hat{\mu} = \exp \left(6,113 - 0,00723X_1 + 0,02180X_2 + 0,00553X_3 - 0,02391X_4 - 0,01672X_5 + 0,00005X_6 - 0,02544X_7 + 1,533X_8 \right)$$

Nilai rasio antara devians dengan derajat kebebasan dari regresi binomial negatif adalah 1,71 merupakan nilai yang cukup mendekati 1 sehingga dapat dikatakan bahwa regresi binomial negatif dapat mengatasi adanya kasus overdispersi pada regresi Poisson.

4.2.3 Pemodelan Jumlah Kejadian Diare di Kota Surabaya Menggunakan Metode GWNBR

Penyakit diare termasuk penyakit yang dapat menular melalui kondisi lingkungan sekitar. Kota Surabaya memiliki perbedaan karakteristik wilayah antara wilayah satu dengan lainnya, misalnya kondisi kepadatan penduduk yang berbeda antar wilayah. Oleh karena itu dapat diduga adanya pengaruh letak geografis suatu wilayah tertentu terhadap jumlah kejadian diare. Sehingga didapatkan antar wilayah akan ada perbedaan faktor-faktor yang mempengaruhi kejadian diare dengan menggunakan analisis GWNBR.

4.2.3.1 Pengujian Dependensi dan Heterogenitas Spasial

Adanya perbedaan karakteristik antara satu titik pengamatan dengan titik pengamatan lainnya dapat dilihat pada pengujian *Breusch-Pagan* dengan pembahasan sebagai berikut.

Hipotesis:

$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_{31}^2 = \sigma^2$ (varians antar lokasi sama)

$H_1 : \text{Minimal ada satu } \sigma_i^2 \neq \sigma^2$ (varians antar lokasi berbeda)

Berdasarkan hasil analisis didapatkan nilai statistik uji *Breusch-Pagan* sebesar 9,523 dengan p-value 0,3001. Selain itu juga diketahui dengan jumlah parameter 8 dan taraf signifikan yang digunakan 5% maka didapatkan $\chi^2_{(8;0,05)}$ sebesar 15,507. Dilihat dari kedua kriteri (statistik uji *Breusch-Pagan* dan p-

value) tersebut maka didapatkan hasil bahwa varians antar lokasi sama.

Pengamatan di suatu lokasi bergantung pada pengamatan di lokasi lain yang letaknya berdekatan dapat diindikasikan adanya dependensi spasial. Pengujian dependensi spasial dapat dilakukan dengan Moran's I dengan pembahasan sebagai berikut.

Hipotesis:

$H_0 : I = 0$ (tidak terdapat dependensi spasial)

$H_1 : I \neq 0$ (terdapat dependensi spasial)

Hasil pengujian Morans'I menunjukkan *p-value* sebesar 0,0017 dengan menggunakan taraf signifikansi 5% maka didapatkan bahwa terdapat dependensi spasial. Hal tersebut menyatakan pengamatan suatu lokasi bergantung pada pengamatan di lokasi lain yang letaknya berdekatan.

Berdasarkan hasil pengujian aspek data spasial yang meliputi pengujian heterogenitas menggunakan *Breusch-Pagan* dan pengujian dependensi menggunakan Morans'I didapatkan hasil bahwa penelitian mengenai jumlah kejadian diare di Surabaya telah memenuhi salah satu diantaranya, yaitu terdapat dependensi spasial. Sehingga selanjutnya dapat dilakukan pemodelan menggunakan metode GWNBR.

Pada pemodelan GWNBR dilakukan dengan menambahkan pembobotan spasial dan dalam penelitian ini menggunakan pembobot fungsi kernel *adaptive bisquare* karena setiap kecamatan memiliki jumlah kejadian diare yang berbeda sehingga *bandwidth* yang dibutuhkan setiap wilayah juga berbeda. Langkah yang perlu dilakukan sebelum menentukan matriks pembobot adalah mencari jarak antar kecamatan yang dapat dilihat pada Lampiran 7. Selanjutnya setelah didapatkan jarak antar kecamatan maka dapat dibentuk matriks pembobot untuk penaksiran parameter di masing-masing kecamatan Kota Surabaya dengan cara memasukkan *bandwidth* dan jarak ke dalam fungsi kernel. Matriks pembobot dapat dilihat pada Lampiran 8.

4.2.3.2 Pengujian Signifikansi Parameter GWNBR

Pengujian signifikansi parameter pada model GWNBR terdiri dari pengujian secara serentak dan parsial. Pengujian serentak dilakukan untuk mengetahui adanya pengaruh atau tidak variabel prediktor terhadap variabel respon secara serentak, dengan hipotesis sebagai berikut.

$$H_0: \beta_1(u_i, v_i) = \beta_2(u_i, v_i) = \dots = \beta_8(u_i, v_i) = 0$$

$$H_1: \text{paling sedikit ada satu } \beta_k(u_i, v_i) \neq 0$$

Berdasarkan hasil analisis didapatkan nilai devians model GWNBR sebesar 18,338. Sedangkan dengan taraf signifikansi 5% diperoleh $\chi^2_{(8;0.05)}$ sebesar 15,507 yang artinya nilai devians lebih besar dari $\chi^2_{(8;0.05)}$. Jadi dapat dikatakan bahwa paling sedikit ada satu parameter model GWNBR yang signifikan berpengaruh maka perlu dilanjutkan dengan pengujian parsial untuk mengetahui variabel yang berpengaruh pada masing-masing kecamatan dengan hipotesis sebagai berikut.

$$H_0: \beta_k(u_i, v_i) = 0$$

$$H_1: \beta_k(u_i, v_i) \neq 0$$

Berdasarkan hasil pengujian diperoleh parameter yang signifikan berbeda-beda untuk tiap kecamatan di Kota Surabaya. Hasil estimasi parameter GWNBR dapat dilihat pada Lampiran 10. Nilai $|Z_{hit}|$ parameter setiap kecamatan dibandingkan dengan nilai $Z_{0,05/2}$ (1,96). Jika nilai $|Z_{hit}| > 1,96$ maka Tolak H_0 , yang artinya parameter signifikan terhadap model. Variabel yang signifikan di setiap kecamatan dapat dilihat pada Tabel 4.6. Berdasarkan hasil menunjukkan bahwa setiap kecamatan memiliki model yang berbeda, namun ada variabel yang berpengaruh di semua kecamatan Kota Surabaya yang disebut dengan variabel global. Variabel global tersebut adalah kepadatan penduduk (X_6), rasio sarana kesehatan puskesmas (X_7) dan rasio dokter umum (X_8).

Tabel 4.6 Penaksiran Parameter Model GWNBR

No.	Kecamatan	Variabel Signifikan
1.	Sukomanunggal	$X_5 X_6 X_7 X_8$
2.	Tandes	$X_3 X_5 X_6 X_7 X_8$
3.	Asemrowo	$X_5 X_6 X_7 X_8$
4.	Benowo	$X_3 X_5 X_6 X_7 X_8$
5.	Pakal	$X_3 X_6 X_7 X_8$
6.	Lakarsantri	$X_3 X_5 X_6 X_7 X_8$
7.	Sambikerep	$X_3 X_5 X_6 X_7 X_8$
8.	Genteng	$X_6 X_7 X_8$
9.	Tegalsari	$X_6 X_7 X_8$
10.	Bubutan	$X_6 X_7 X_8$
11.	Simokerto	$X_5 X_6 X_7 X_8$
12.	Pabean Cantikan	$X_2 X_4 X_5 X_6 X_7 X_8$
13.	Semampir	$X_2 X_5 X_6 X_7 X_8$
14.	Krembangan	$X_2 X_4 X_5 X_6 X_7 X_8$
15.	Bulak	$X_5 X_6 X_7 X_8$
16.	Kenjeran	$X_2 X_5 X_6 X_7 X_8$
17.	Tambaksari	$X_5 X_6 X_7 X_8$
18.	Gubeng	$X_6 X_7 X_8$
19.	Rungkut	$X_5 X_6 X_7 X_8$
20.	Tenggilis Mejoyo	$X_5 X_6 X_7 X_8$
21.	Gunung Anyar	$X_5 X_6 X_7 X_8$
22.	Sukolilo	$X_2 X_5 X_6 X_7 X_8$
23.	Mulyorejo	$X_5 X_6 X_7 X_8$
24.	Sawahan	$X_6 X_7 X_8$
25.	Wonokromo	$X_6 X_7 X_8$
26.	Karang Pilang	$X_5 X_6 X_7 X_8$
27.	Dukuh Pakis	$X_3 X_5 X_6 X_7 X_8$
28.	Wiyung	$X_5 X_6 X_7 X_8$
29.	Gayungan	$X_6 X_7 X_8$
30.	Wonocolo	$X_6 X_7 X_8$
31.	Jambangan	$X_6 X_7 X_8$

Berdasarkan Tabel 4.6 didapatkan pengelompokkan kecamatan dengan melihat kesamaan variabel yang signifikan. Faktor-faktor yang berpengaruh terhadap kejadian diare di Kota

Surabaya adalah persentase RT memiliki jamban sehat (X_2), persentase RT memiliki tempat sampah sehat (X_3), persentase RT miskin (X_4), persentase RT ber-PHBS (X_5), kepadatan penduduk (X_6), rasio sarana kesehatan (X_7) dan rasio dokter umum (X_8). Berikut ini disajikan Tabel 4.7 mengenai enam kelompok kecamatan dengan kesamaan variabel yang signifikan.

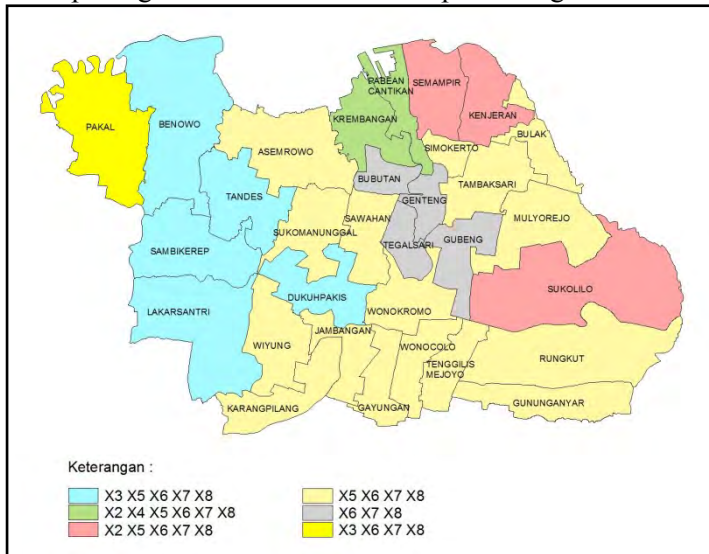
Tabel 4.7 Pengelompokan Kecamatan Berdasarkan Kesamaan Variabel Signifikan dalam Model GWNBR

Kel.	Kecamatan	Variabel Signifikan
1	Tandes, Benowo, Dukuh Pakis, Lakarsantri, Sambikerep	$X_3 X_5 X_6 X_7 X_8$
2	Pabean Cantikan, Krembangan	$X_2 X_4 X_5 X_6 X_7 X_8$
3	Semampir, Kenjeran, Sukolilo	$X_2 X_5 X_6 X_7 X_8$
4	Sukomanunggal, Asemrowo, Simokerto, Bulak, Tambaksari, Rungkut, Tenggilis Mejoyo, Gunung Anyar, Mulyorejo, Karangpilang, Wiyung, Sawahan, Wonokromo, Gayungan, Wonocolo, Jambangan	$X_5 X_6 X_7 X_8$
5	Genteng, Tegalsari, Bubutan, Gubeng,	$X_6 X_7 X_8$
6	Pakal	$X_3 X_6 X_7 X_8$

Berdasarkan Tabel 4.7 dapat diketahui bahwa ada variabel global, yaitu variabel yang berpengaruh terhadap kejadian diare pada semua kecamatan di Kota Surabaya. Variabel global tersebut adalah kepadatan penduduk (X_6), rasio sarana kesehatan puskesmas (X_7) dan rasio dokter umum (X_8). Semakin padat penduduk akan menyebabkan lingkungan yang kumuh sehingga memicu kejadian diare lebih meningkat sedangkan semakin banyak tersedia sarana kesehatan puskesmas maka bisa memberikan fasilitas untuk penyuluhan pencegahan penyakit diare sehingga dapat mengurangi kejadian diare. Jika rasio dokter umum meningkat kejadian diare juga meningkat, hal ini kemungkinan dikarenakan sudah tersedia dokter umum namun

masyarakat menganggap remeh penyakit diare adalah penyakit ringan sehingga tidak melakukan pengobatan ke dokter.

Agar lebih jelas mengenai kelompok pemetaan kecamatan maka dapat digambarkan dalam bentuk peta sebagai berikut.



Gambar 4.10 Pemetaan Kecamatan Berdasarkan Faktor-Faktor Mempengaruhi Jumlah Kejadian Diare di Kota Surabaya

Berdasarkan Gambar 4.10 diketahui bahwa terdapat 6 kelompok dengan wilayah berdekatan, yang cenderung memiliki karakteristik sama. Persentase rumah tangga memiliki jamban sehat (X_2) berpengaruh terhadap kejadian diare di lima kecamatan, sedangkan persentase rumah tangga memiliki tempat sampah sehat berpengaruh terhadap kejadian diare di enam kecamatan dan untuk persentase rumah tangga miskin berpengaruh terhadap kejadian diare hanya di dua kecamatan. Namun untuk persentase rumah tangga ber-PHBS berpengaruh terhadap kejadian diare di 26 kecamatan. Jadi berdasarkan delapan variabel yang diduga mempengaruhi kejadian diare, menunjukkan hanya ada satu variabel yang tidak berpengaruh di

tiap kecamatan Kota Surabaya yaitu persentase rumah sehat (X_1). Hal ini dikarenakan kemungkinan rumah tangga sudah tergolong rumah sehat namun tidak diimbangi dengan pola hidup bersih dan sehat yang menyebabkan keadaan rumah sehat tidak berarti dan tidak menimbulkan pengaruh terhadap kejadian diare.

Selain itu pada Gambar 4.10 juga menunjukkan pemetaan kecamatan yang sesuai dengan kondisi nyata di Kota Surabaya. Surabaya Pusat yang ditunjukkan oleh warna abu-abu memiliki variabel berpengaruh hanya variabel global yang secara umum juga berpengaruh di kecamatan lainnya. Hal tersebut sesuai dengan keadaan di pusat Kota Surabaya yang telah menunjukkan kebersihan lingkungan, misalnya tersedia tempat sampah sehat, RT ber-PHBS dan sebagainya. Namun jika bergeser ke arah timur dan utara yang berwarna kuning variabel yang berpengaruh selain variabel global juga ditambah dengan persentase RT ber-PHBS, hal tersebut nampak dikarenakan daerah kecamatan yang cenderung berdekatan dengan pantai dan memiliki kepadatan penduduk yang cukup tinggi sehingga masih dijumpai RT yang belum menerapkan hidup ber-PHBS. Begitu juga dengan Kecamatan Semampir, Kenjeran dan Sukolilo dimana wilayahnya masih ada RT yang tidak memiliki jamban sehat sehingga variabel tersebut juga berpengaruh terhadap kejadian diare di ketiga kecamatan tersebut.

Semakin bergeser ke utara juga diketahui ada penambahan variabel yang berpengaruh yaitu persentase RT miskin, hal tersebut dikarenakan di Kecamatan Krembangan dan Pabean Cantikan masih tergolong cukup tinggi untuk hal persentase RT miskin. Namun jika semakin bergeser ke arah barat ditambahkan dengan variabel yang berpengaruh adalah persentase RT memiliki tempat sampah sehat. Berdasarkan hal tersebut diketahui bahwa kondisi di sekitar kecamatan Surabaya Barat masih ada yang tertinggal akan kepemilikan tempat sampah sehat. Sehingga hal tersebut berpengaruh terhadap sanitasi lingkungan yang menyebabkan kejadian diare.

Sebagai contoh akan disajikan lokasi penelitian ke-16 (u_{16}, v_{16}) yaitu Kecamatan Kenjeran sebagai berikut.

Tabel 4.8 Parameter Model GWNBR di Kecamatan Kenjeran

Parameter	Estimasi	Z
$\hat{\beta}_0$	6,113015	22534100*
$\hat{\beta}_1$	0,000912	0,090024
$\hat{\beta}_2$	0,032453	2,055251*
$\hat{\beta}_3$	-0,0111	-0,75978
$\hat{\beta}_4$	-0,03619	-1,81795
$\hat{\beta}_5$	-0,01716	-2,56355*
$\hat{\beta}_6$	0,00004	1729,839*
$\hat{\beta}_7$	-0,02491	-2025,05*
$\hat{\beta}_8$	1,533431	590930,2*

*) signifikan dengan taraf nyata 5%

Variabel yang memberikan pengaruh dapat dilihat dari nilai $|Z|$ yang lebih besar dari $z_{(\alpha/2)}$, dengan taraf signifikansi 5% didapatkan 1,96 sehingga dapat diketahui bahwa di Kecamatan Kenjeran variabel yang berpengaruh terhadap jumlah kejadian diare adalah persentase RT memiliki jamban sehat (X_2), persentase RT ber-PHBS (X_5), kepadatan penduduk (X_6), rasio sarana kesehatan puskesmas (X_7) dan rasio dokter umum (X_8). Sehingga didapatkan bentuk model sebagai berikut.

$$\hat{\mu} = \exp \left(\begin{array}{l} 6,113015 + 0,000912X_1 + 0,032453X_2 - 0,0111X_3 - 0,03619X_4 \\ -0,01716X_5 + 0,00004X_6 - 0,2491X_7 + 1,533431X_8 \end{array} \right)$$

Berdasarkan model yang terbentuk di Kecamatan Kenjeran maka diketahui bahwa setiap pertambahan satu persen RT berperilaku hidup bersih dan sehat akan mengurangi kejadian diare sebesar $\exp(0,01716) = 1,0173 \approx 1$ kejadian dengan asumsi variabel lain konstan. Hal ini sesuai dengan semakin banyak RT yang ber-PHBS maka akan terbentuk sanitasi lingkungan yang baik sehingga akan mengurangi rata-rata kejadian penyakit diare karena virus dapat ditularkan pada lingkungan yang tidak sehat.

Setiap pertambahan satu jiwa per km^2 akan menambah rata-rata kejadian diare sebesar $\exp(0,00004) = 1,00004 \approx 1$ kejadian dengan asumsi variabel lain konstan. Semakin padat jumlah

penduduk maka akan didapatkan peluang lingkungan kumuh semakin besar sehingga penularan penyakit melalui virus dapat tersebar dengan cepat. Sarana kesehatan masyarakat juga menunjukkan variabel yang berpengaruh sehingga diketahui bahwa setiap kenaikan satu satuan rasio sarana kesehatan puskesmas maka akan mengurangi rata-rata jumlah kejadian diare sebesar $\exp(0,2491) = 1,2828 \approx 1$ kejadian dengan asumsi variabel lain konstan. Hal ini sesuai karena peranan puskesmas di berbagai kalangan masyarakat tergolong penting. Jadi dengan adanya sarana untuk kesehatan maka akan mengurangi adanya penyakit yang diderita oleh masyarakat terutama untuk penyakit diare.

Berdasarkan variabel yang berpengaruh di Kecamatan Kenjeran didapatkan prediksi kejadian diare sebesar 4061 kejadian. Untuk lebih jelas mengenai data prediksi kejadian diare berdasarkan faktor-faktor yang mempengaruhi di setiap kecamatan dapat dilihat pada Lampiran 12.

4.2.3.3 Pemilihan Model Terbaik

Pemilihan model terbaik digunakan kriteria AIC pada ketiga model yang dilakukan, yaitu model regresi Poisson, regresi Binomial Negatif dan GWNBR dengan hasil sebagai berikut.

Tabel 4.9 Hasil Perhitungan Kriteria AIC

Metode Analisis	AIC
Regresi Poisson	9111,1
Regresi Binomial Negatif	511,49
GWNBR	187,903

Berdasarkan Tabel 4.9 menunjukkan bahwa dari ketiga model tersebut, model GWNBR memiliki nilai AIC paling kecil dibandingkan dengan regresi Poisson dan regresi Binomial Negatif. Jadi dapat disimpulkan bahwa model GWNBR lebih baik dalam memodelkan jumlah kejadian diare di Kota Surabaya.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil dan pembahasan yang telah dilakukan, maka dapat disimpulkan sebagai berikut.

1. Rata-rata kejadian diare di Kota Surabaya adalah sebanyak 2093 kejadian dengan kejadian diare paling sedikit sebanyak 467 kejadian di Kecamatan Pakal dengan kepadatan penduduk 2189,61 dan kejadian diare paling banyak adalah Kecamatan Sawahan sebanyak 5657 kejadian dengan kepadatan penduduk sebesar 20516,06. Di Kota Surabaya persentase rumah sehat, persentase rumah tangga memiliki jamban sehat dan persentase rumah tangga memiliki tempat sampah sehat sudah ada beberapa kecamatan yang mencapai maksimum yaitu 100 persen. Sedangkan rata-rata persentase rumah tangga ber-PHBS adalah 69,81 persen. Kepadatan penduduk di Kota Surabaya paling tinggi adalah Kecamatan Simokerto yaitu sebesar 31194 jiwa per km². Dalam upaya pembangunan kesehatan Kota Surabaya memiliki rata-rata rasio dokter umum sebesar 0,668 dan sarana kesehatan puskesmas sebesar 0,238.
2. Berdasarkan analisis pemodelan GWNBR didapatkan variabel global yang berpengaruh terhadap kejadian diare pada 31 kecamatan di Kota Surabaya yaitu kepadatan penduduk (X_6), rasio sarana kesehatan puskesmas (X_7) dan rasio dokter umum (X_8). Pemodelan menggunakan GWNBR menghasilkan model yang berbeda tiap kecamatan sehingga terbentuk enam kelompok model dengan kesamaan variabel yang berpengaruh terhadap kejadian diare. Persentase rumah tangga memiliki jamban sehat (X_2) berpengaruh terhadap kejadian diare di lima kecamatan, sedangkan persentase rumah tangga memiliki tempat sampah sehat berpengaruh terhadap kejadian diare di enam kecamatan dan untuk persentase rumah tangga miskin berpengaruh terhadap

kejadian diare hanya di dua kecamatan. Namun untuk persentase rumah tangga ber-PHBS berpengaruh terhadap kejadian diare di 26 kecamatan.

5.2 Saran

Saran yang didapatkan berdasarkan hasil penelitian adalah dengan adanya variabel global yang berpengaruh terhadap kejadian diare maka dapat mendeteksi kejadian diare di kecamatan Kota Surabaya. Jadi untuk pencegahan penyakit diare dapat melalui penyuluhan yang diadakan di kecamatan yang padat penduduknya agar masyarakat mengerti akan pentingnya sanitasi lingkungan. Selain itu juga diberi masukan agar ketika mengalami diare dapat berobat di dokter umum yang tersedia di puskesmas masing-masing kecamatan sehingga fasilitas yang disediakan oleh pemerintah sebagai upaya pengendalian penyakit diare dapat dimaksimalkan dengan baik.

DAFTAR PUSTAKA

- Agresti, A. 2002. *Categorical Data Analysis Second Edition*. New York: John Wiley & Sons.
- Andjarsari, N. I. 2012. *Analisis Faktor-Faktor yang Mempengaruhi Terjangkitnya Penyakit Diare di Provinsi Nanggroe Aceh Darussalam dengan Menggunakan Regresi Logistik Biner*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Anselin, L. 1998. *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer Academic Publishers.
- Arumsari, Nurvita. 2011. *Kajian Geographically Weighted Lasso (GWL) Untuk Pemodelan Penderita Diare Di Kabupaten Sumedang Dalam Pencapaian Target Open Defecation Free (ODF)*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Barus, B., dan Wiradisastara, U. S., 2000. *Sistem Informasi Geografi: Sarana Manajemen Sumberdaya*. Bogor: Institut Pertanian Bogor.
- Cameron, A. C., & Trivedi, P. K. 1998. *Regression Analysis of Count Data*. Cambridge: Camb Cambridge University.
- Dinas Kesehatan, K. S. 2013. *Profil Kesehatan Kota Surabaya*. Surabaya: Dinas Kesehatan Kota Surabaya.
- Dinas Kesehatan, P. J. 2013. *Profil Kesehatan Provinsi Jawa Timur*. Surabaya: Dinas Kesehatan Provinsi Jawa Timur.
- Famoye, F., Wulu, J., & Singh, K. 2004. On The Generalize Poisson Regression Model with an Application to Accident Data. *Journal of Data Science* 2, 287-295.
- Greene, W. 2008. *Functional Forms for the Negative Binomial Model for Count Data, Foundation, and Trends in Econometrics*, 99, 585-590. New York: New York University.
- Hilbe, J. 2011. *Negative Binomial Regression, Second Edition*. New York: Cambridge University Press.

- Hocking, R. 1996. *Method and Applications of Linier Models*. New York: Cambridge University Press.
- Hosmer, D. W., Lemeshow, S., Sturdivant, R. X. 2013. *Applied Logistic Regression Third Edition*. United States of America: John Wiley and Sons Inc.
- Juniardi, L. C., & Salamah M. 2015. *Analisis Faktor-Faktor yang mempengaruhi Jumlah Kasus Kusta di Jawa Timur pada Tahun 2013 Menggunakan Geographically Weighted Negative Binomial Regression (GWNBR)*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Mc Cullagh, P., & Nelder, J. 1989. *Generalized Linear Models Second Edition*. London: Chapman & Hall.
- Nandasari, B. N. 2014. *Pemodelan Jumlah Kejadian Luar Biasa Difteri di Jawa Timur Menggunakan Geographically Weighted Negative Binomial Regression (GWNBR)*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Pratama, Wahendra. 2015. *Pemetaan dan Pemodelan Jumlah Kasus Tuberculosis (TBC) di Provinsi Jawa Barat dengan Pendekatan Geographically Weighted Negative Binomial Regression (GWNBR)*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Purwaningsih, H. 2009. *Analisis Hubungan Antara Kondisi Sanitasi, Air Bersih dan Penderita Diare di Jawa Timur*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Ricardo, A., & Carvalho, T. 2013. *Geographically Weighted Negative Binomial Regression-Incorporating Overdispersion*. Business Media New York: Springer Science.
- Silva A.F. 2008. *Faktor sanitasi yang berpengaruh pada penderita diare di Kabupaten Sidoarjo*. Surabaya: Universitas Wijaya Kusuma.
- Walpole, E. R. 1995. *Pengantar Statistik Edisi Ketiga*. Jakarta: Pustaka Utama.
- Wikipedia. 2015. *Penyakit Diare*. <http://id.wikipedia.org/wiki/Diare> (diakses pada 20 April 2015, 19.10 WIB).

LAMPIRAN

Lampiran 1. Data Jumlah Kejadian Diare Beserta Faktor-Faktor yang diduga Mempengaruhinya

No	Kecamatan	Y	X ₁	X ₂	X ₃	X ₄	X ₅
1	Sukomanunggal	3457	71.03	99.70	97.02	6.8	79.03
2	Tandes	3476	85.72	98.71	97.66	7.39	65.87
3	Asemworo	968	70.97	59.89	96.51	9.19	46.65
4	Benowo	1155	55.43	98.87	100	6.68	84.70
5	Pakal	467	87.86	99.38	41.54	5.07	84.40
6	Lakarsantri	2066	74.68	99.82	75.16	10.28	65.43
7	Sambikerep	1219	90.39	100	86.10	9.4	87.06
8	Genteng	3157	87.90	95.68	74.55	13.56	55.12
9	Tegalsari	2308	72.92	92.86	66.98	13.74	88.60
10	Bubutan	1872	72.13	98.99	87.93	12.52	79.33
11	Simokerto	2256	77.29	98.98	99.63	24.41	66.76
12	Pabean Cantikan	1323	67	99.61	69.51	23.03	51.07
13	Semampir	2007	80.01	94.01	76.28	28	63.40
14	Krembangan	2416	73.52	97.02	99.83	13.17	55.77
15	Bulak	984	74.15	94.33	87.56	8.44	71.05
16	Kenjeran	5602	74.62	99.26	90.34	8.57	59.64
17	Tambaksari	2096	87.3	99.87	98.62	12.87	63.09
18	Gubeng	2485	78.83	98.76	85.8	5.83	83.22
19	Rungkut	2101	92.71	99.30	96.97	10.13	79.87
20	Tenggilis	1173	82.90	99.24	81.05	4.87	86.75
21	Gunung Anyar	842	100	96.27	100	9.75	91.61
22	Sukolilo	3106	80.8	99.18	74.74	10.66	54.89
23	Mulyorejo	1512	89.86	99.05	95.38	7.17	83.69
24	Sawahan	5657	82.19	97.34	85.06	11.98	65.36
25	Wonokromo	2388	83.65	96.80	94.38	10.12	34.25
26	Karangpilang	1471	90.94	96.31	35.35	7.26	58.03
27	Dukuh Pakis	659	78	95.96	76.95	10.41	73.30
28	Wiyung	1524	76.70	99.96	88.83	6.98	68.38
29	Gayungan	1110	96	98.86	97.98	5.79	88.33
30	Wonocolo	2407	92.85	98.31	98.10	6.53	48.47
31	Jambangan	1612	80.28	98.52	100	9.35	80.78

Lampiran 1. Data Jumlah Kejadian Diare Beserta Faktor-Faktor yang diduga Mempengaruhinya (lanjutan)

Kecamatan	X₆	X₇	X₈	U	V
Sukomanunggal	11102.49	0.5855	0.2439	7.27	112.699
Tandes	9484.19	0.5714	0.2381	7.257	112.673
Asemworo	2816.77	0.6897	0.2299	7.239	112.689
Benowo	2322.49	0.5441	0.2267	7.229	112.647
Pakal	2189.61	0.6213	0.2071	7.232	112.614
Lakarsantri	2934.5	1.4461	0.3615	7.323	112.653
Sambikerep	3326.19	1.1506	0.2876	7.275	112.655
Genteng	11735.89	1.2654	0.5272	7.26	112.743
Tegalsari	20324.71	0.6881	0.2007	7.274	112.737
Bubutan	22287.82	0.6974	0.2034	7.247	112.729
Simokerto	31193.82	0.7426	0.2475	7.24	112.753
Pabean Cantikan	10397.94	0.4242	0.2121	7.222	112.733
Semampir	17605.94	0.5835	0.1783	7.215	112.749
Krembangan	13025.9	0.0826	0.0275	7.227	112.724
Bulak	6534.83	0.7915	0.1978	7.232	112.789
Kenjeran	19310.67	0.7209	0.2853	7.216	112.772
Tambaksari	23203.67	0.4314	0.1198	7.251	112.769
Gubeng	16334.42	0.4597	0.1532	7.265	112.752
Rungkut	5850.57	0.4865	0.2027	7.317	112.802
Tenggilis	13370.65	0.4064	0.1354	7.324	112.756
Gunung Anyar	6515.86	0.4741	0.1975	7.339	112.789
Sukolilo	5153.69	0.7371	0.2252	7.291	112.804
Mulyorejo	13923.67	0.6218	0.2072	7.269	112.793
Sawahan	20516.06	0.6905	0.3021	7.274	112.721
Wonokromo	9548.42	0.6633	0.2948	7.294	112.738
Karangpilang	7996.53	0.4064	0.2032	7.332	112.686
Dukuh Pakis	6583.6	0.4584	0.1910	7.292	112.7
Wiyung	5557.62	0.8664	0.3249	7.309	112.693
Gayungan	7168.53	0.6894	0.2872	7.328	112.727
Wonocolo	12059.59	1.1007	0.3669	7.325	112.744
Jambangan	11286.16	0.6343	0.3171	7.324	112.716

Predictor	Coef	SE Coef	T	P	VIF
Constant	945	3266	0.29	0.775	
x1	-17.52	20.85	-0.84	0.410	1.121
x2	31.81	29.14	1.09	0.287	1.219
x3	4.01	12.07	0.33	0.743	1.082
x4	-71.98	43.23	-1.66	0.110	1.594
x5	-29.77	14.73	-2.02	0.056	1.357
x6	0.10085	0.03256	3.10	0.005	1.547
x7	-615	1250	-0.49	0.628	3.327
x8	5249	3991	1.32	0.202	3.572

Lampiran 4. Hasil Analisis Regresi Poisson

Deviance Residuals:

Min	1Q	Median	3Q	Max
-22.570	-14.718	-7.207	11.251	33.794

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	7.048e+00	9.016e-02	78.171	< 2e-16 ***
x1	-1.182e-02	4.890e-04	-24.179	< 2e-16 ***
x2	1.920e-02	8.133e-04	23.607	< 2e-16 ***
x3	1.361e-03	2.865e-04	4.753	2.01e-06 ***
x4	-3.619e-02	9.593e-04	-37.728	< 2e-16 ***
x5	-1.443e-02	3.070e-04	-46.990	< 2e-16 ***
x6	4.743e-05	6.896e-07	68.785	< 2e-16 ***
x7	-3.443e-01	2.883e-02	-11.941	< 2e-16 ***
x8	2.434e+00	8.387e-02	29.016	< 2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

Null deviance: 19422.0 on 30 degrees of freedom
 Residual deviance: 8803.8 on 22 degrees of freedom
 AIC: 9111.1

Number of Fisher Scoring iterations: 4

Lampiran 5. Hasil Analisis Regresi Binomial Negatif

Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.6803	-1.0007	-0.3510	0.6616	1.6440

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	6.113e+00	1.163e+00	5.257	1.47e-07 ***
x1	-7.234e-03	7.418e-03	-0.975	0.32948
x2	2.180e-02	1.037e-02	2.102	0.03556 *
x3	5.528e-03	4.295e-03	1.287	0.19805
x4	-2.391e-02	1.537e-02	-1.555	0.11987
x5	-1.672e-02	5.236e-03	-3.192	0.00141 **
x6	4.598e-05	1.158e-05	3.972	7.14e-05 ***
x7	-2.544e-02	4.444e-01	-0.057	0.95436
x8	1.533e+00	1.419e+00	1.081	0.27984

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Negative Binomial(7.5189) family taken to be 1)

Null deviance: 71.272 on 30 degrees of freedom
 Residual deviance: 31.671 on 22 degrees of freedom
 AIC: 511.49

Number of Fisher Scoring iterations: 1

Theta: 7.52
 Std. Err.: 1.88

2 x log-likelihood: -491.491

Lampiran 6. Hasil Analisis Pengujian Aspek Data Spasial**HETEROGENITAS - BP TEST**

studentized Breusch-Pagan test

data: modelnegbin

BP = 9.523, df = 8, p-value = 0.3001

DEPENDENSI - MORAN'S I

\$observed

[1] 0.1735548

\$expected

[1] -0.03333333

\$sd

[1] 0.06624587

\$p.value

[1] 0.001789971

Lampiran 7. Jarak Antar Kecamatan

	1	2	3	4	5	...	28	29	30	31
1	0.000	0.029	0.033	0.066	0.093	...	0.039	0.064	0.071	0.057
2	0.029	0.000	0.024	0.038	0.064	...	0.056	0.089	0.098	0.080
3	0.033	0.024	0.000	0.043	0.075	...	0.070	0.097	0.102	0.089
4	0.066	0.038	0.043	0.000	0.033	...	0.092	0.127	0.136	0.117
5	0.093	0.064	0.075	0.033	0.000	...	0.110	0.148	0.160	0.137
6	0.070	0.069	0.091	0.094	0.099	...	0.042	0.074	0.091	0.063
7	0.044	0.025	0.050	0.047	0.059	...	0.051	0.089	0.102	0.078
8	0.045	0.070	0.058	0.101	0.132	...	0.070	0.070	0.065	0.069
9	0.038	0.066	0.059	0.101	0.130	...	0.056	0.055	0.051	0.054
10	0.038	0.057	0.041	0.084	0.116	...	0.072	0.081	0.079	0.078
11	0.062	0.082	0.064	0.107	0.139	...	0.091	0.092	0.085	0.092
12	0.059	0.069	0.047	0.086	0.119	...	0.096	0.106	0.104	0.103
13	0.074	0.087	0.065	0.103	0.136	...	0.109	0.115	0.110	0.114
14	0.050	0.059	0.037	0.077	0.110	...	0.088	0.101	0.100	0.097
15	0.098	0.119	0.100	0.142	0.175	...	0.123	0.114	0.103	0.117
16	0.091	0.107	0.086	0.126	0.159	...	0.122	0.121	0.113	0.122
17	0.073	0.096	0.081	0.124	0.156	...	0.096	0.088	0.078	0.090
18	0.053	0.079	0.068	0.111	0.142	...	0.074	0.068	0.061	0.069
19	0.113	0.142	0.137	0.178	0.206	...	0.109	0.076	0.059	0.086
20	0.079	0.107	0.108	0.145	0.169	...	0.065	0.029	0.012	0.040
21	0.113	0.142	0.141	0.180	0.205	...	0.101	0.063	0.047	0.075
22	0.107	0.135	0.126	0.169	0.199	...	0.112	0.085	0.069	0.094
23	0.094	0.121	0.108	0.151	0.183	...	0.108	0.089	0.074	0.095
24	0.022	0.051	0.047	0.087	0.115	...	0.045	0.054	0.056	0.050
25	0.046	0.075	0.074	0.112	0.139	...	0.047	0.036	0.032	0.037
26	0.063	0.076	0.093	0.110	0.123	...	0.024	0.041	0.058	0.031
27	0.022	0.044	0.054	0.082	0.105	...	0.018	0.045	0.055	0.036
28	0.039	0.056	0.070	0.092	0.110	...	0.000	0.039	0.053	0.027
29	0.064	0.089	0.097	0.127	0.148	...	0.039	0.000	0.017	0.012
30	0.071	0.098	0.102	0.136	0.160	...	0.053	0.017	0.000	0.028
31	0.057	0.080	0.089	0.117	0.137	...	0.027	0.012	0.028	0.000

Lampiran 8. Matriks Pembobot Geografis

	1	2	3	4	5	...	28	29	30	31
1	1.000	0.718	0.653	0.043	0.000	...	0.516	0.062	0.007	0.176
2	0.826	1.000	0.879	0.709	0.309	...	0.442	0.020	0.000	0.099
3	0.770	0.871	1.000	0.616	0.119	...	0.187	0.000	0.000	0.007
4	0.522	0.824	0.778	1.000	0.866	...	0.212	0.000	0.000	0.016
5	0.415	0.692	0.589	0.912	1.000	...	0.251	0.010	0.000	0.051
6	0.502	0.516	0.256	0.225	0.176	...	0.799	0.455	0.260	0.585
7	0.709	0.899	0.645	0.680	0.513	...	0.626	0.128	0.026	0.258
8	0.342	0.000	0.099	0.000	0.000	...	0.000	0.000	0.019	0.000
9	0.465	0.002	0.053	0.000	0.000	...	0.097	0.117	0.179	0.129
10	0.612	0.257	0.557	0.000	0.000	...	0.047	0.000	0.002	0.005
11	0.295	0.040	0.260	0.000	0.000	...	0.000	0.000	0.016	0.000
12	0.458	0.301	0.627	0.092	0.000	...	0.020	0.000	0.000	0.000
13	0.296	0.143	0.430	0.016	0.000	...	0.000	0.000	0.000	0.000
14	0.567	0.423	0.745	0.166	0.000	...	0.054	0.000	0.000	0.003
15	0.072	0.000	0.053	0.000	0.000	...	0.000	0.000	0.033	0.000
16	0.188	0.045	0.241	0.000	0.000	...	0.000	0.000	0.017	0.000
17	0.125	0.000	0.038	0.000	0.000	...	0.000	0.003	0.063	0.000
18	0.207	0.000	0.012	0.000	0.000	...	0.000	0.014	0.088	0.007
19	0.004	0.000	0.000	0.000	0.000	...	0.016	0.336	0.562	0.208
20	0.176	0.000	0.000	0.000	0.000	...	0.366	0.845	0.973	0.721
21	0.054	0.000	0.000	0.000	0.000	...	0.157	0.583	0.753	0.447
22	0.000	0.000	0.000	0.000	0.000	...	0.000	0.132	0.342	0.053
23	0.001	0.000	0.000	0.000	0.000	...	0.000	0.021	0.157	0.001
24	0.817	0.252	0.323	0.000	0.000	...	0.377	0.188	0.159	0.266
25	0.376	0.000	0.000	0.000	0.000	...	0.343	0.585	0.666	0.555
26	0.492	0.324	0.127	0.010	0.000	...	0.916	0.764	0.557	0.862
27	0.862	0.507	0.322	0.000	0.000	...	0.903	0.492	0.307	0.658
28	0.689	0.438	0.215	0.005	0.000	...	1.000	0.696	0.474	0.842
29	0.257	0.003	0.000	0.000	0.000	...	0.672	1.000	0.930	0.968
30	0.245	0.001	0.000	0.000	0.000	...	0.510	0.941	1.000	0.849
31	0.406	0.080	0.010	0.000	0.000	...	0.837	0.969	0.830	1.000

Lampiran 9. Z Hitung Pengujian Parsial GWNBR

	β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8
1	439293.6	-0.02849	0.976719	1.733116	-0.82717	-2.20518	2238.48	-2422.07	1570638
2	917510.7	-0.27774	1.521314	3.565258	-0.9054	-2.41445	2553.182	-1273.7	46465775
3	1884654	0.045803	1.726867	1.37942	-1.41453	-2.422	2250.038	-3565.11	3725036
4	1886464	-0.48727	1.220069	5.384295	-0.81748	-1.9778	1798.658	-1160.2	56115312
5	1747945	-0.46297	1.157157	5.409558	-0.53637	-1.88878	1435.914	-437.339	8308563
6	215571.1	-1.08984	1.642565	2.155924	-0.06106	-2.18951	1645.74	-16.5834	53260.14
7	446388.9	-0.57344	1.672828	4.625112	-0.40729	-2.53033	2248.592	-69.23	315167.9
8	10329609	-0.00623	0.705313	-0.31974	-1.02231	-1.11018	912.2193	-1818.08	633356.7
9	2477526	-0.28302	0.724643	-0.51363	-0.70503	-1.79531	1200.885	-915.898	654024
10	7393368	0.07466	1.556725	0.044455	-1.64766	-1.53056	1361.422	-1812.77	611836
11	21224184	-0.02063	1.841468	-0.61085	-1.8304	-2.3322	1643.267	-2140.89	638128.4
12	14228719	0.212234	2.00366	-0.1479	-1.98257	-2.06944	1744.199	-2325.64	723076.2
13	20282822	0.144014	1.965311	-0.33572	-1.9092	-2.14037	1650.113	-1797.82	554529.6
14	10517381	0.271884	2.025331	-0.0632	-1.98293	-2.11365	1855.87	-2740.86	891707.8
15	8837273	0.044045	1.888276	-1.16499	-1.54908	-2.76353	1628.64	-2250.29	533519
16	22534100	0.090024	2.055251	-0.75978	-1.81795	-2.56355	1729.839	-2025.05	590930.2
17	8660839	-0.16456	1.745452	-1.28043	-1.50847	-2.80425	1589.019	-2443.92	551838.7
18	3980611	-0.28903	0.855598	-0.90903	-0.94504	-1.71357	1003.498	-2447.86	445455.3
19	4001136	-0.26174	1.856361	-0.41155	0.170862	-2.4374	158.179	-882.291	462252.8
20	3698594	-0.87091	1.702849	-0.32489	0.217787	-3.89807	845.2998	-10714.5	757194.6
21	3985873	-0.63797	1.875567	-0.40041	0.25592	-3.05511	308.3958	-17060.7	2666888
22	3450875	-0.19111	1.990435	-0.61648	-0.0506	-2.26609	371.4426	-2089.54	458401.8
23	2854197	-0.05948	1.92153	-1.08405	-0.58027	-2.54672	927.1963	-2909.05	595772.1
24	971153.4	-0.2895	1.016786	0.700784	-0.85596	-2.37594	1852.736	-2115.13	1110743
25	1639494	-0.67706	0.860003	-0.35409	-0.30358	-3.48767	1416.817	-289.078	52789.02
26	366095.3	-1.3087	1.696453	0.894505	-0.05701	-2.70356	1765.681	-35.0884	66437.11
27	320065.7	-0.66407	1.344053	2.134439	-0.35981	-2.75195	2260.307	-66.3872	117011.1
28	304057.4	-1.23728	1.776262	1.549769	-0.16679	-2.88587	2148.163	-40.2371	75785.29
29	2016358	-0.72307	1.015814	0.244122	0.008455	-3.47045	1348.34	-780.799	137451.9
30	3233013	-0.86723	1.43495	-0.03332	0.151832	-3.94921	1150.438	-3824.27	430439
31	1906089	-1.15836	1.409358	0.167961	-0.08743	-3.55467	1682.936	-109.942	64979.53

Lampiran 10. Koefisien Parameter

	θ	β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8
1	7.519	6.112	0.000	0.014	0.009	-0.078	-0.015	0.00008	-0.026	1.533
2	7.519	6.113	-0.002	0.018	0.010	-0.066	-0.020	0.00008	-0.026	1.533
3	7.519	6.113	0.000	0.019	0.006	-0.053	-0.019	0.00006	-0.025	1.533
4	7.519	6.113	-0.003	0.018	0.012	-0.054	-0.023	0.00007	-0.026	1.533
5	7.519	6.113	-0.003	0.019	0.013	-0.053	-0.025	0.00007	-0.026	1.533
6	7.519	6.113	-0.011	0.022	0.005	-0.012	-0.016	0.00006	-0.024	1.533
7	7.519	6.113	-0.004	0.019	0.010	-0.056	-0.020	0.00007	-0.027	1.533
8	7.519	6.113	0.000	0.021	-0.003	-0.037	-0.008	0.00004	-0.025	1.533
9	7.519	6.113	-0.007	0.029	-0.005	-0.053	-0.012	0.00006	-0.026	1.533
10	7.519	6.113	0.001	0.018	0.000	-0.038	-0.009	0.00004	-0.025	1.533
11	7.519	6.113	0.000	0.024	-0.004	-0.035	-0.012	0.00004	-0.025	1.533
12	7.519	6.113	0.003	0.019	-0.001	-0.037	-0.012	0.00004	-0.025	1.533
13	7.519	6.113	0.002	0.021	-0.002	-0.035	-0.014	0.00004	-0.025	1.533
14	7.519	6.113	0.003	0.019	0.000	-0.039	-0.012	0.00004	-0.025	1.533
15	7.519	6.113	0.001	0.032	-0.011	-0.036	-0.017	0.00004	-0.025	1.533
16	7.519	6.113	0.002	0.026	-0.005	-0.035	-0.016	0.00004	-0.025	1.533
17	7.519	6.113	-0.003	0.034	-0.011	-0.037	-0.015	0.00004	-0.025	1.533
18	7.519	6.113	-0.007	0.036	-0.012	-0.042	-0.013	0.00005	-0.025	1.533
19	7.519	6.113	-0.015	0.039	-0.010	0.020	-0.011	0.00001	-0.025	1.533
20	7.519	6.113	-0.015	0.028	-0.002	0.022	-0.011	0.00003	-0.025	1.533
21	7.519	6.113	-0.018	0.034	-0.004	0.029	-0.011	0.00001	-0.025	1.533
22	7.519	6.113	-0.009	0.039	-0.013	-0.004	-0.011	0.00002	-0.025	1.533
23	7.519	6.113	-0.002	0.038	-0.017	-0.027	-0.014	0.00003	-0.025	1.533
24	7.519	6.113	-0.004	0.018	0.004	-0.057	-0.012	0.00007	-0.026	1.533
25	7.519	6.113	-0.014	0.032	-0.004	-0.044	-0.014	0.00007	-0.026	1.533
26	7.519	6.113	-0.015	0.025	0.002	-0.009	-0.013	0.00006	-0.024	1.533
27	7.519	6.113	-0.008	0.019	0.006	-0.050	-0.013	0.00008	-0.027	1.533
28	7.519	6.113	-0.013	0.024	0.003	-0.024	-0.013	0.00007	-0.025	1.533
29	7.519	6.113	-0.010	0.018	0.001	0.001	-0.012	0.00007	-0.025	1.533
30	7.519	6.113	-0.012	0.023	0.000	0.017	-0.011	0.00004	-0.025	1.533
31	7.519	6.113	-0.015	0.025	0.000	-0.014	-0.013	0.00007	-0.025	1.533

Lampiran 11. *Syntax* Analisis Menggunakan *Software R*

```
#REGRESI POISSON#
library(MASS)
modelpoisson=glm(y~x1+x2+x3+x4+x5+x6+x7+x8,family=poisson,data
a=data)
summary(modelpoisson)
#REGRESI BINOMIAL NEGATIF#
modelnegbin<-glm.nb(y~x1+x2+x3+x4+x5+x6+x7+x8,data=data)
summary(modelnegbin) #RBP

#BPTEST#
library(lmtest)
depen=lm(y~x1+x2+x3+x4+x5+x6+x7+x8, data=data)
bptest(depen)
bptest(modelnegbin) #RBP

library(spgwr)
bdwtBisquare=ggwr.sel(y~x1+x2+x3+x4+x5+x6+x7+x8,data=data,coor
ds=cbind(data$u,data$v),adapt=TRUE,gweight=gwr.bisquare)
GRTGB=ggwr(y~x1+x2+x3+x4+x5+x6+x7+x8,data=data,coords=cbind
(data$u,data$v),adapt=bdwtBisquare,gweight=gwr.bisquare)
GRTGB$bandwidth
#JARAK#
u=data[,10]
u<-as.matrix(u)
i<-nrow(u)
v=data[,11]
v<-as.matrix(v)
j<-nrow(v)
library(fields)
jarak<-matrix(nrow=31,ncol=31)
for(i in 1:31)
for(j in 1:31){jarak[i,j]=sqrt((u[i,]-u[j,])**2+(v[i,]-v[j,])**2)}
write.table(jarak,file="D:/jarak1.csv",sep="," )
```

Lampiran 11. *Syntax Analisis Menggunakan Software R*
(Lanjutan)

```
#PEMBOBOT#
bdwtBisquare<- GRTGB$bandwidth
bdwtBisquare<- as.matrix(bdwtBisquare)
bdwtBisquare
i<-nrow(bdwtBisquare)
pembobotB<-matrix(nrow=31,ncol=31)
for(i in 1:31)
for(j in 1:31)
{pembobotB[i,j]=(1-(jarak[i,j]/bdwtBisquare[i,])**2)**2
pembobotB[i,j]<-
ifelse(jarak[i,j]<bdwtBisquare[i,],pembobotB[i,j],0)}
write.table(pembobotB,file="D:/pembobotB1.csv",sep=",")
#MORAN'S I#

library(MASS)
gemes2=function(X,y,W1,phi1,b1){
beta=matrix(c(0),20,10,byrow=T)
#beta[1,1]=phi1
#beta[1,2:10]=c(b1)
satu<-rep(1,31)
satu<-as.matrix(satu)
bxx <- c(b1)
b01<-rbind(c(phi1,c(b1)))
for(i in 1:20){
Xb1<-as.matrix(X)%*%as.matrix(bxx)
mul<-exp(Xb1)
delta11<-((log(1+phi1*mul)-
digamma(y+(1/phi1))+digamma(1/phi1))/phi1^2)+((y-
mul)/((1+phi1*mul)*phi1))
```

Lampiran 11. *Syntax Analisis Menggunakan Software R*
(Lanjutan)

```

delta11<-as.matrix(delta11)
p11<-t(satu)%*%W1)%*%delta11
delta21<-(y-mu1)/(1+phi1*mu1)
delta21<-as.matrix(delta21)
p21<-t(X)%*%as.matrix(W1)%*%delta21
p21<-as.matrix(p21)
gt1<-rbind(p11,p21)
delta31<-((trigamma(y+(1/phi1))-
trigamma(1/phi1))/phi1^4)+((2*digamma(y+(1/phi1))-
2*digamma(1/phi1)-
2*log(1+phi1*mu1))/phi1^3)+((2*mu1)/(phi1^2*(1+phi1*mu1)))
+(((y+(1/phi1))*mu1^2)/(1+phi1*mu1)^2)-(y/phi1^2)
delta31<-as.matrix(delta31)
p31<-t(satu)%*%W1)%*%delta31
p31<-as.matrix(p31)
delta41<-mu1*(mu1-y)/(1+phi1*mu1)^2
delta41<-as.matrix(delta41)
p41<-t(X)%*%W1)%*%delta41
p41<-as.matrix(p41)
h11<-rbind(p31,p41)
delta51<-mu1*(phi1*y+1)/(1+phi1*mu1)^2
delta51<-t(delta51)
delta51<-c(delta51)
delta51<-as.matrix(diag(delta51))
p51<-t(X)%*%as.matrix(W1)%*%delta51)%*%as.matrix(X)
p51<--1*p51
p51<-as.matrix(p51)
h21<-rbind(t(p41),p51)
H1<-cbind(h11,h21)
H11<-ginv(H1)
beta[i,]<-(t(b01)-H11)%*%gt1)
phi1<-beta[i,1] #RBP
b01<-t(beta[i,]) #RBP

```


Lampiran 11. *Syntax* Analisis Menggunakan *Software R* (Lanjutan)

```

bxx<-beta[i,2:10] #RBP
}
return(list(beta=beta,hessian=H1))
}
gwnbr1 <- function(x,y,W,teta){
# beta <- ginv(t(x) %*% x) %*% t(x) %*% y
teta<-modelnegbin$theta
beta<- as.matrix(modelnegbin$coefficients) #RBP
param <- matrix(c(0),nrow(x),ncol(x)+1, byrow=T)
zhit <- matrix(c(0),nrow(x),ncol(x), byrow=T)
for(i in 1:31){
ww <- as.matrix(diag(W[i,]))
hit <- gemes2(x,y,ww,teta,beta)
parameter<- hit$beta #RBP
param[i,] <- hit$beta[20,]
write.csv(hit$hessian,file=paste("hessian",i,".csv"))
invh <- -ginv(as.matrix(hit$hessian))
for(j in 1:ncol(x)){
zhit[i,j] <- param[i,j+1] / invh[j+1,j+1]
}
}
return(list(koefisien=param,Z_hitung=zhit,parameter=parameter))
#RBP
}
#data=read.csv("e:/bismillah.csv",header=TRUE, sep=",")
#bobot=as.matrix(read.csv("D://pembobotB1.csv",
header=FALSE, sep=","))
bobot=pembobotB #RBP
xx=data[,2:9]
y=data[,1]
x=as.matrix(cbind(1,xx))
phx1<-modelnegbin$theta #RBP
mod=gwnbr1(x,y,bobot,phx1) #RBP

```

Lampiran 11. *Syntax Analisis Menggunakan Software R*
(Lanjutan)

```

mod$Z_hitung
mod$koefisien
mod$parameter
write.csv(mod$koefisien,file="D:/koefisien.csv")
write.csv(mod$Z_hitung,file="D:/Z_hitung.csv")
datay<-as.matrix(data[,1])
datax <- as.matrix(cbind(1,data[,2:9]))
tetanb<- modelnegbin$theta
betanb<- as.matrix(modelnegbin$coefficients)
muw<-as.matrix(rep(exp(betanb[1]),31))
slr<-matrix(0, nrow(data),1)
for(i in 1:nrow(data)){
slr[i]<-0
for(r in 1:datay[i]){slr[i]<-slr[i]+log(r+(1/tetanb))}
}
Lw<-sum(slr-lgamma(datay+1)+datay*log(tetanb*muw)-
(datay+(1/tetanb))*log(1+tetanb*muw))
muo<-exp(datax%%*%betanb)
Lo<-sum(slr-lgamma(datay+1)+datay*log(tetanb*muo)-
(datay+(1/tetanb))*log(1+tetanb*muo))
DNB<-2*(Lo-Lw)
DNB
tetagw<-as.matrix(mod$koefisien[,1])
betagw<-as.matrix(mod$koefisien[,2:10])
muwgw<- as.matrix(exp(mod$koefisien[,2]))
muogw<-as.matrix(exp(apply(datax*betagw,1,sum)))
Lwgw<-sum(slr-lgamma(datay+1)+datay*log(tetagw*muogw)-
(datay+(1/tetagw))*log(1+tetagw*muogw))
Logw<-sum(slr-lgamma(datay+1)+datay*log(tetagw*muogw)-
(datay+(1/tetagw))*log(1+tetagw*muogw))
DGW<-2*(Logw-Lwgw)
DGW

```

Lampiran 12. Prediksi Kejadian Diare di Surabaya Berdasarkan Model GWNBR

Kecamatan	Prediksi Kejadian Diare
Sukomanunggal	2714
Tandes	2841
Asemworo	1093
Benowo	1257
Pakal	575
Lakarsantri	1684
Sambikerep	961
Genteng	3439
Tegalsari	2117
Bubutan	2857
Simokerto	3026
Pabean Cantikan	1705
Semampir	1441
Krembangan	1864
Bulak	1489
Kenjeran	4061
Tambaksari	2668
Gubeng	2426
Rungkut	1491
Tenggilis	1442
Gunung Anyar	986
Sukolilo	2914
Mulyorejo	1776
Sawahan	3873
Wonokromo	2518
Karangpilang	1390
Dukuh Pakis	1136
Wiyung	1904
Gayungan	1049
Wonocolo	2557
Jambangan	1906

BIODATA PENULIS



Penulis bernama lengkap Fitri Ernawati yang biasa dipanggil Erna, lahir di Surabaya pada tanggal 07 April 1992. Penulis merupakan anak kelima dari lima bersaudara. Penulis menyelesaikan pendidikan di TK Among Putro Surabaya 1998, SDN Gading IX Surabaya tahun 2004, SMPN 9 Surabaya tahun 2007, SMAN 1 Surabaya tahun 2010 dan penulis melanjutkan studinya ke Perguruan Tinggi. Penulis diterima di Jurusan Statistika ITS pada tahun 2010 melalui jalur seleksi masuk program studi DIII. Penulis lulus pada tahun 2013 yang kemudian dilanjutkan studi Lintas Jalur (LJ) melalui seleksi masuk program studi LJ S1 di Jurusan Statistika ITS.

Selama di bangku perkuliahan penulis aktif dalam berorganisasi, diantaranya menjadi staf HRD Divisi PSt HIMASTA-ITS 2011/2012, staf keuangan BEM FMIPA ITS 2011/2012. Penulis juga pernah tergabung dalam kegiatan UKM (Unit Kegiatan Mahasiswa) olahraga basket. Selain itu penulis juga memiliki pengalaman bekerja sebagai surveyor di PT. BHSTEX, MPM, YAMAHA, Enciety Bussiness Consult dan sebagainya, serta sebagai Data Entri dan Data Analisis di Regional Economic Development Institute (REDI), HYDYNAMIX, LPPM ITS. Prinsip penulis dalam hidupnya adalah *“Don’t Give Up Before Trying”*. Untuk kritik dan saran yang membangun atau ingin berdiskusi lebih lanjut mengenai Tugas Akhir ini dapat menghubungi penulis melalui email fitrierna07@gmail.com